Durham University

Durham E-Theses

Potential-based Formulations of the Navier-Stokes Equations and their Application

FLORIAN MINKUS MARNER

How to cite:

MARNER, FLORIAN MINKUS (2019) Potential-based Formulations of the Navier-Stokes Equations and their Application. Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a https://etheses.durham.ac.uk/id/eprint/13270/ is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
https://etheses.durham.ac.uk

Potential-based Formulations of the Navier-Stokes Equations and their Application

by
Florian Minkus Marner

Submitted in partial fulfilment
of the requirements for the degree
Doctor of Philosophy



Department of Engineering

February 2019

Abstract

Based on a Clebsch-like velocity representation and a combination of classical variational principles for the special cases of ideal and Stokes flow a novel discontinuous Lagrangian is constructed; it bypasses the known problems associated with non-physical solutions and recovers the classical Navier-Stokes equations together with the balance of inner energy in the limit when an emerging characteristic frequency parameter tends to infinity. Additionally, a generalized Clebsch transformation for viscous flow is established for the first time. Next, an exact first integral of the unsteady, three-dimensional, incompressible Navier-Stokes equations is derived; following which gauge freedoms are explored leading to favourable reductions in the complexity of the equation set and number of unknowns, enabling a self-adjoint variational principle for steady viscous flow to be constructed. Concurrently, appropriate commonly occurring physical and auxiliary boundary conditions are prescribed, including establishment of a first integral for the dynamic boundary condition at a free surface. Starting from this new formulation, three classical flow problems are considered, the results obtained being in total agreement with solutions in the open literature.

A new least-squares finite element method based on the first integral of the steady two-dimensional, incompressible, Navier-Stokes equations is developed, with optimal convergence rates established theoretically. The method is analysed comprehensively, thoroughly validated and shown to be competitive when compared to a corresponding, standard, primitive-variable, finite element formulation. Implementation details are provided, and the well-known problem of mass conservation addressed and resolved via selective weighting. The attractive positive definiteness of the resulting linear systems enables employment of a customized scalable algebraic multigrid method for efficient error reduction. The solution of several engineering related problems from the fields of lubrication and film flow demonstrate the flexibility and efficiency of the proposed method, including the case of unsteady flow, while revealing new physical insights of interest in their own right.

Publications

During the course of the research undertaken leading to the presentation of this thesis, the following publications appeared in the open literature:

F. Marner, P. H. Gaskell and M. Scholle: On a potential-velocity formulation of Navier-Stokes equations. *Physical Mesomechanics* **17** (4) (2014) 341-348.

M. Scholle and F. Marner: A generalized Clebsch transformation leading to a first integral of Navier–Stokes equations. *Physics Letters A* **380** (40) (2016) 3258–3261.

M. Scholle and F. Marner: A non-conventional discontinuous Lagrangian for viscous flow. *Royal Society Open Science* **4** (2) (2017) 160447.

F. Marner, P. H. Gaskell and M. Scholle: A complex-valued first integral of Navier-Stokes equations: Unsteady Couette flow in a corrugated channel system. *Journal of Mathematical Physics* **58** (4) (2017) 043102.

M. Scholle, P. H. Gaskell and F. Marner: Exact integration of the unsteady incompressible Navier-Stokes equations, gauge criteria, and applications. *Journal of Mathematical Physics* **59** (4) (2018) 043101.

F. Marner, M. Scholle, D. Herrmann and P. H. Gaskell: Competing Lagrangians for incompressible and compressible viscous flow. *Royal Society Open Science* **6** (1) (2019) 181595.

M. Scholle, P. H. Gaskell and F. Marner: A potential field description for gravity-driven film flow over piece-wise planar topography. *Fluids* **4** (2) (2019) 4020082.

Acknowledgements

Undertaking the research contained in this thesis has been a life-shaping experience that would not have been possible without the support and guidance I received from many people.

I would like start by saying a very big thank you to both my supervisors: Prof. Markus Scholle, Heilbronn University, and Prof. Philip H. Gaskell, Durham University, for all the support and encouragement they gave me. It is not only their profound scientific knowledge in the fields of theoretical physics, fluid mechanics and numerical methods, their extraordinary creativity and diversity of ideas but also their continuous engagement, way beyond the call of duty, that enabled a thesis in its present form to be achieved. Vivid scientific debates and frequent helpful feedback facilitated progress, accompanied by discussions of British and German politics, history, literature, sports and philosophy over lunch. Joint conference participation across Europe, time in Durham, Phil's visits to Heilbronn, and all the time we spent together beyond "work" will remain a lasting and cherished memory. Thanks are due also to Dr. Sergii Veremieiev, Durham University, for sharing his detailed knowledge of and helpful advice in the fields of film flow, Computational Fluid Dynamics and numerical algorithms.

I gratefully acknowledge the funding received in support of my studies from the Deutsche Forschungsgemeinschaft (DFG), covering a good portion of my time at Heilbronn and enabling me to participate in several inspiring scientific conferences. Furthermore, I greatly appreciate the support, by the Deutscher Akademischer Austauschdienst (DAAD), of Phil's semester in Heilbronn, and of my tuition fees at Durham University by the Thomas Gessmann-Stiftung; as to the latter, special thanks go to Gerhard Häußermann from the foundation and Prof. emer. Rudolf Kern for valuable advice and establishing the contact. For enduring support and encouragement, when bridging a funding gap, thanks go to our former Dean, Prof. Oliver Lenzen, his assistant, Dr. Elena Dickert, and our Head of Department, Mechatronics and Robotics, Prof. Peter Ott, as well as all the other members of the department. My brief time as Head of the Control Engineering Laboratory and collaborations with Prof. Ipek Saraç Heinz and my colleague Jan Neumeister was a fruitful period, expanding my research scope and experience in many ways.

Among the many Bachelor and Masters project students I had the pleasure of being involved with, I would like to mention in particular Marcel Mellmann and Delphine Her-

rmann who contributed, with extraordinary motivation and thoroughness, a great deal to the detailed elaboration, coding and testing of the spectral film flow method, as well as related topics.

Particular thanks go to my colleagues Jens Gerdes, Georg Wörle and Timo Hufnagel who made settling in at Heilbronn University as comfortable as possible, introducing me to the daily university routine and all its peculiarities; manifold support and vivid discussions throughout the whole of my time in Heilbronn made arriving at the office each day a pleasure. In the same vein, thanks go to departmental secretaries Christa Bücking and Waltraud Bayer for practical tips on how to survive the administrative system and for always having time to listen. Deserving of special mention are the many room-mates and neighbours I encountered during my time in Heilbronn: Daniel Uzelmaier, Günther Grimm, Jan Neumeister, Joachim Walter, Anton Fogel and Marc Diesse. The frequent and varied discussions of engineering and mathematics problems, Latex typesetting, lecture preparation, politics, sports etc. made my time in Heilbronn an incredibly fun one. Thanks also to Jens Klaski for helping me improve my table soccer skills through regular "training sessions" while at the same time discussing boulder tricks and his injection moulding experiments. Last but not least, thanks to all the people attending the regular "Weißwurstfrühstück" and coffee breaks, "institutionalised" by Jens, during which a very enjoyable atmosphere prevailed with an active exchange of opinions and ideas, providing light relief and a welcome distraction.

Finally, I wish to express my deep and sincere gratitude to the members of my family. My parents, Bruni and Manfred, are undeniably responsible for the critically thinking and scientifically oriented person I have become – without your exceptional and continual stimulation together with your unfailing support, in every respect, I would have lacked the courage to embark on the research reported in this thesis and the endurance to accomplish it. Thanks also to my brothers, Miklas and Marcel, for the great times we regularly spend together – your enduring encouragement has seen me through the tough times. Last but not least I owe a huge debt of gratitude to Mariele for her invaluable contribution to the success of my work – your never-ending patience, encouragement and understanding accompanied me throughout the entire journey.

Declaration

The author confirms that the work submitted in this thesis is his own and that appropriate credit has been given where reference has been made to the work of others. No part of the thesis has been submitted elsewhere in respect of the award of any other degree or qualification.

Copyright ©: F. M. Marner 2019

The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.

Contents

1 Introduction **1**
1.1 Maxwell's vector potential and the advent of gauge theories 3
1.2 Potentials in fluid dynamics and the limits of Lagrange formalism 10
1.2.1 Potentials and first integrals of the equations of motion 10
1.2.2 Potentials and variational principles 17
1.3 Objectives and thesis structure 30
1.3.1 Part I 30
1.3.2 Part II 32

I New Potential Field Representations of the Navier-Stokes Equations 35

2 Approaches based on Clebsch-like representation **37**
2.1 Construction of the Lagrangian 38
2.1.1 Conventional approach and examples 38
2.1.2 Constructing a non-conventional Lagrangian based on complex fields 45
2.1.3 Incompressible flow with constant specific heat and external force . . 48
2.2 Variation of the discontinuous Lagrangian: general formalism 48
2.2.1 Euler-Lagrange equations 49
2.2.2 Matching conditions 50
2.2.3 Production condition 51
2.3 Resulting equations of motion and matching conditions 53
2.3.1 Equations of motion 53
2.3.2 Matching conditions 54
2.3.3 Production condition and thermodynamic aspects 55
2.4 Discussion 56
2.5 A generalized Clebsch transformation 58
2.5.1 Extension to viscous flow 59
2.5.2 Axisymmetric stagnation flow as an example 61

3 Approaches based on Goursat-like representation **65**
3.1 Overview 65
3.2 Two-dimensional unsteady and incompressible flow 67
3.2.1 Formulation of the field equations 68
3.2.2 Formulation of boundary conditions 69
3.2.3 Application of the method to simple flows 72
3.3 General three-dimensional unsteady and incompressible flow 75
3.3.1 Preliminaries and introduction of a streamfunction vector 76

3.3.2 First integral of the field equations . . . 76
3.4 Closure via selective gauge criteria . . . 78
3.4.1 Convenient re-ordering of the first integral . . . 79
3.4.2 Traceless form . . . 81
3.4.3 Self-adjoint form . . . 82
3.5 Application of the Methodology . . . 85
3.5.1 Boundary conditions . . . 85
3.5.2 Unsteady stagnation flow . . . 87
3.5.3 Flow within a cubic domain . . . 90
3.5.4 Steady Stokes flow . . . 96
3.5.5 Steady free surface flow in the presence of symmetries . . . 98

II Construction of a New and Efficient Potential-based FE Solver – Implementation and Application 103

4 Discrete formulation and method of solution 105
4.1 Preliminaries . . . 106
4.2 First order system formulation . . . 108
4.3 Weak form and convergence analysis of the Stokes problem . . . 111
4.3.1 A priori estimates for the linear equations . . . 111
4.3.2 Weak variational formulation . . . 117
4.3.3 Error of the least-squares method . . . 123
4.3.4 Least-squares FEM in irregular domains . . . 128
4.4 Analysis of the non-linear problem . . . 133
4.4.1 Non-linear variational problem . . . 133
4.4.2 Discretisation error estimates . . . 135

5 Implementation and solution of linear systems 141
5.1 Method of solution . . . 141
5.1.1 Structure of the fully non-linear problem . . . 141
5.1.2 Treatment of periodic problems . . . 146
5.1.3 Isoparametric concept . . . 149
5.2 Conditioning of linear systems . . . 161
5.3 Numerical tests and validation . . . 163
5.3.1 Linear test cases and results . . . 164
5.3.2 Non-linear test cases and results . . . 173
5.4 Algebraic multigrid . . . 176
5.4.1 Scalar AMG: basic ideas . . . 177
5.4.2 AMG as a preconditioner . . . 183
5.4.3 AMG and the first integral approach . . . 184
5.4.4 Validation of the AMG approach . . . 188

6 Engineering applications 201
6.1 Friction reduction between lubricated surfaces . . . 201
6.1.1 Problem description and solution approach . . . 201

6.1.2 Mathematical modelling . . . 203
6.1.3 Lubrication solution . . . 205
6.1.4 A parametric study . . . 207
6.2 Gravity-driven steady film flow over periodic topography . . . 217
6.2.1 Spectral complex-valued solution representation . . . 218
6.2.2 Boundary conditions in terms of analytic functions . . . 220
6.2.3 Fourier series discretisation . . . 222
6.2.4 Spectral iterative method of solution . . . 225
6.2.5 Alternative solution by the LSFEM . . . 227
6.2.6 Numerical results . . . 229
6.3 Film flow over non-planar surfaces . . . 236
6.3.1 Steady axisymmetric flow over a hemisphere . . . 236
6.4 Unsteady Couette flow confined between two corrugated rigid surfaces . . . 240
6.4.1 Equations of motion . . . 241
6.4.2 Asymptotic model and method of solution . . . 242
6.4.3 Numerical model and method of solution . . . 245
6.4.4 Results and discussion . . . 247

7 Conclusions and future work 257
7.1 Summary and discussion of results . . . 259
7.2 Suggestions for future work . . . 268

A Viscous flow and the 3D first integral 273
A.1 Variational principle for viscous flow - Calculations . . . 273
A.1.1 Symmetries and associated Noether balances of (2.4) . . . 273
A.1.2 Steady solution of (2.8-2.10) for Poiseuille flow . . . 274
A.1.3 Euler-Lagrange equations of Lagrangian (2.32) . . . 276
A.1.4 Equations of motion . . . 276
A.1.5 Derivation of the production condition (2.42) using distributions . . . 277
A.2 Three-dimensional first integral . . . 279
A.2.1 Recovery of the 2D form as a special case . . . 279
A.2.2 Free surface boundary conditions . . . 279
A.2.3 First integral with non-planar surfaces . . . 282
A.2.4 Finite difference scheme . . . 284

B Functional analysis and elliptic PDE theory 289
B.1 Brief compendium of functional analysis . . . 289
B.1.1 Banach and Hilbert spaces . . . 289
B.1.2 Sobolev spaces . . . 290
B.1.3 Traces and continuation . . . 294
B.1.4 Dual spaces . . . 296
B.1.5 Functional derivatives . . . 298
B.2 Review of Agmon-Douglis-Nirenberg theory . . . 299
B.3 Auxiliary problem and regularity considerations . . . 304
B.4 Auxiliary problem in an infinite sector . . . 308

C Numerical method: further underpinnings **317**
C.1 Convergence results: Taylor-Couette flow . . . 317
C.2 Convergence results: Colliding flow . . . 320
C.3 Taylor-Hood discretisation of the NS equations . . . 320
C.4 Scalar algebraic multigrid: the details . . . 324
C.4.1 Algebraic smoothing . . . 324
C.4.2 Coarsening procedure . . . 327
C.4.3 Interpolation . . . 332
C.4.4 Convergence of the two-level cycle . . . 335

D Application related calculations **339**
D.1 Film flow over periodically corrugated topography . . . 339
D.1.1 Calculations for the dynamic BC (6.26) . . . 339
D.1.2 Calculations for the dynamic BC (6.28) . . . 340
D.2 Film flow over non-planar surfaces . . . 342
D.2.1 Boundary conditions for flow over a hemisphere . . . 342
D.2.2 Comparison with heuristic considerations . . . 343
D.3 Turnstile lobe effect . . . 344
D.3.1 General considerations for small Reynolds numbers . . . 344
D.3.2 Fourier discretization . . . 345

Bibliography **347**

1 Introduction

For many systems of differential equations arising in science and engineering the introduction of auxiliary potential fields has proved beneficial in several respects. Maxwell's equations serve as a prominent example, in which the electric field and the magnetic flux density can be favourably replaced by a scalar and a vector potential, respectively. They enable a reduction in the number of variables and hence equations, while appropriate usage of the gauge freedoms facilitates their reformulation as two decoupled wave equations leading to a self-adjoint system with an associated Lagrangian [133]. Apart from the obvious formulaic advantages, the potentials allow for the treatment of electrodynamics in the framework of Hamilton's principle. In this context it is arguable that the historical debate surrounding the Maxwell potentials, dating back to the 19th century [134, 283], can be considered formative for the usage of potential fields in modern theoretical physics and ground-breaking with respect to the later development of gauge theory [141], suggesting Maxwell theory as a significant reference point for other field theories.

The present work explores and contributes to the utilisation of potentials in the field of *fluid mechanics* with Maxwell theory guiding the objectives. The use of potentials that are simply defined as those fields which lead to certain degrees of gauge freedom and therefore are physically non-measurable, represents a vast area dating back to the work of Lagrange [82] and Clebsch [64] at the end of the 18th and middle of the 19th centuries, respectively. The review provided subsequently places the present work in context, showing why the introduction of potentials in the case of viscous fluid flow in particular has remained rather limited and what can be done to (partly) resolve this issue.

In fluid mechanics in general auxiliary potentials are frequently used either to reduce the number of unknowns, as in the case of defining a streamfunction [154], or to express physical quantities in terms of more tractable functions. The latter is achieved for instance by the Papkovich-Neuber representation of velocity or the complex-valued Goursat representation of the streamfunction, depending only on harmonic and analytic functions, respectively [111, 154]. Most of these theoretically useful representations are limited either with respect to the spatial dimension, symmetry requirements or the generality of the physics involved.

As a usefully important application, potential fields have been employed to enable the *integration* of the governing equations of fluid motion. Most notably, Bernoulli's equation

is obtained as a first integral of Euler's equations in the absence of vorticity and viscosity, if the velocity vector is taken to be the gradient of a scalar potential. The so-called Clebsch transformation [154, 191] and related approaches allow for further extension to flows with non-vanishing vorticity, resulting in a generalised Bernoulli equation complemented with transport equations for the Clebsch potentials [236]. The term "first integral" in this specific context was, amongst others, shaped by Helmholtz [123] in his treatise on vortex motion, where he also introduced what is today called Helmholtz-Hodge decomposition of vector fields into an irrotational and a solenoidal part. Most generally, the task is to reformulate the equations as a gradient of a scalar quantity or the divergence of a tensorial quantity set to zero allowing for integration in a further step; this may be achieved using auxiliary potential fields and does *not* necessarily imply a reduction of the total differential order.

While the above examples refer to inviscid flow in particular, progress involving the full incompressible Navier-Stokes (NS) equations has been far less fruitful and restricted in the main to the limiting case of steady two-dimensional (2D) flow, the most recent contribution prior to the work reported in this thesis being that of Scholle et al. [230]. Along similar lines to Legendre [159], Coleman [70] and Ranger [199], they constructed an exact complex-valued first integral, based on the introduction of an auxiliary complex field; the formulation embodies the classical complex-valued Goursat representation for steady Stokes flow [8, 111], allowing the streamfunction to be expressed in terms of two analytic functions [171].

Closely related to the integration of field equations is the establishment of *variational principles*. For physical systems formulated within the framework of Lagrange's formalism the dynamics are completely defined by one function only: the Lagrangian. This methodical concept successfully applies, for example, to the conservative branch of Newtonian mechanics. Contrary to this, in continuum theories many open problems remain unsolved to this day, especially when considering dissipative processes [216]; the viscous flow of a fluid, described by the NS equations, is a typical example where a Lagrangian formulation is missing [99]. However, several approaches exist for the cases of ideal [99, 183, 213, 236] and Stokes flow [124] which can serve as a starting point.

Certainly, much of the work in this thesis is influenced by the idea of integrating the viscous flow equations encapsulating at least some of the advantages gained in the case of pure Eulerian flow. Closely related to the investigation of potential-based velocity representations is the search for a variational principle of viscous flow. Prior to addressing the specific content of the thesis, the following section provides a comprehensive overview of the important literature that informed the direction of travel and underpins the outcomes.

1.1 Maxwell's vector potential and the advent of gauge theories

Any review of the significance of potential representations in modern physics would be diminished considerably without reference to Maxwell's theory of classical electrodynamics and its immense impact on nearly all major 20th century theoretical developments such as field theory in general, special relativity, wave mechanics and even gauge theory in the modern description of interactions via the standard model of particle physics. Certainly, the influence of Maxwell's ideas on subsequent developments with all their subtleties is one of the great narratives of science, told by numerous authors [134, 283, 284], culminating in the famous words of Richard Feynman [96]:

> "*From a long view of the history of mankind — seen from, say, ten thousand years from now — there can be little doubt that the most significant event of the nineteenth century will be judged as Maxwell's discovery of the laws of electrodynamics. The American Civil War will pale into provincial insignificance in comparison with this important scientific event of the same decade.*"

However, there is one aspect in the history of classical electrodynamics which is particularly fruitful in the context of the present work, namely the treatment of so-called potential fields, associated with which is the question of the most adequate representation of a theory.

Between 1855 and 1865, in a series of three papers, cf. [283] and references therein, James Clerk Maxwell integrated the apparently disparate laws and phenomena of electricity and magnetism, as observed by his contemporaries Charles-Augustin de Coulomb, Carl Friedrich Gauss, André-Marie Ampère, Michael Faraday and others, into a unified theory by modifying Ampère's circuital law and introducing a displacement current term. He had thus distilled an entire body of work on electricity and magnetism into a few equations only and showed, as a consequence, that light propagates as electromagnetic waves; thus, electricity, magnetism and light — three phenomena that seem so vastly different in their manifestation — appeared as different aspects of one and the same phenomenon.

The form of the classical Maxwell equations in terms of the observables, the electric field $\vec{E}$ and the magnetic field $\vec{B}$, which has been taught in universities through most of the 20th century and stated in most standard textbooks is due to Oliver Heaviside (1885) and Heinrich Hertz (between 1884 and 1892); Maxwell's preference was for one in terms of potential fields [283]. There has been a long-standing debate on the use and physicality of those potential fields which can be considered ground-breaking and which turned out to be essential in the later description of gauge theory. In the following, a closer look is taken at the different descriptions of classical electrodynamics.

In a vacuum the partial differential equations (PDEs) for the electric field E_i and the

magnetic flux density B_i in the Heaviside-Hertz form are[1]:

$$\partial_i B_i = 0\,, \tag{1.1a}$$

$$\partial_t B_i + \varepsilon_{ijk}\partial_j E_k = 0\,, \tag{1.1b}$$

$$\epsilon_0 \partial_i E_i = \rho\,, \tag{1.1c}$$

$$\varepsilon_{ijk}\partial_j B_k - \epsilon_0\mu_0\partial_t E_i = \mu_0 j_i\,, \tag{1.1d}$$

for a given charge density ρ and current density j_i. The above set of equations is reduced by introducing a vector potential $\vec{A}$ according to:

$$B_i = \varepsilon_{ijk}\partial_j A_k\,. \tag{1.2}$$

Inserting this into (1.1b) leads to $\varepsilon_{ijk}\partial_j[\partial_t A_k + E_k] = 0$, which on the introduction of the scalar potential φ, such that $\partial_t A_k + E_k = -\partial_k\varphi$, leads to the expression:

$$E_i = -\,\partial_i\varphi - \partial_t A_i\,, \tag{1.3}$$

for the electric field. Hence, by making use of the scalar potential φ and the vector potential A_i, the homogeneous Maxwell equations (1.1a), (1.1b) are fulfilled identically whereas the inhomogeneous ones (1.1c), (1.1d) become PDEs of second order in terms of φ and A_i which are endowed with certain beneficial features:

(i) Written in terms of φ and A_i, equations (1.1c), (1.1d) become self-adjoint [133], that is they can be derived from a variational principle based on the Lagrangian:

$$\ell = \frac{1}{2\mu_0}\left(\varepsilon_{ijk}\partial_j A_k\right)^2 - \frac{\epsilon_0}{2}\left(\partial_i\varphi + \partial_t A_i\right)^2 + \rho\varphi + j_i A_i\,; \tag{1.4}$$

whereas the original Maxwell equations are not self-adjoint.

(ii) If χ is an arbitrary scalar field, then by means of the gauge transformations:

$$A_i \to A_i' = A_i + \partial_i\chi\,, \tag{1.5}$$

$$\varphi \to \varphi' = \varphi - \partial_t\chi\,, \tag{1.6}$$

A_i and φ can be redefined such that the observables B_i and E_i in (1.2), (1.3) remain invariant. Thus a gauge of the potentials, by the fixing of χ, can be employed in such a way as to achieve a convenient form of the remaining field equations.

One of the most common gauge transformations is the *Lorenz gauge*; by a proper choice

[1] Throughout the thesis both tensor calculus and the more classical "nabla calculus" are employed as convenient. Usually the Einstein summation convention is assumed for repeated indices in a mathematical expression.

of χ the identity:

$$\epsilon_0\mu_0\partial_t\varphi + \partial_i A_i = 0\,, \tag{1.7}$$

is fulfilled, leading to the equations (1.1c), (1.1d) having the following decoupled form:

$$\left\{-\frac{1}{c^2}\partial_t^2 + \partial_k^2\right\}\varphi = -\frac{\rho}{\epsilon_0}\,, \tag{1.8a}$$

$$\left\{-\frac{1}{c^2}\partial_t^2 + \partial_k^2\right\}A_i = -\frac{j_i}{\epsilon_0}\,, \tag{1.8b}$$

of a known type, namely two d'Alembert equations, where $c = 1/\sqrt{\epsilon_0\mu_0}$ is the speed of light. For completeness it is noted that all of the above outcomes can similarly be arrived at via the *Coulomb gauge* [133]:

$$\partial_i A_i = 0\,, \tag{1.9}$$

in place of (1.7) resulting in a Poisson equation for φ in place of the inhomogeneous d'Alembert equation (1.8a).

Furthermore, Maxwell's equations can be expressed in a covariant form [133], that is the laws of classical electromagnetism written in a form that is manifestly invariant under Lorentz transformations, in the formalism of special relativity using rectilinear inertial coordinate systems. These expressions make it both simple to prove that the laws of classical electromagnetism take the same form in any inertial coordinate system and provide a way to translate the fields and forces from one frame to another. Combining A_k with φ leads to the 4-potential A^μ while combining E_i with B_i yields an electromagnetic field tensor $F^{\mu\nu}$, the forms of which are:

$$(A^\mu) = \begin{pmatrix} \varphi/c, \\ A_1 \\ A_2 \\ A_3 \end{pmatrix}\,, \qquad (F^{\mu\nu}) = \begin{pmatrix} 0 & -E_1/c & -E_2/c & -E_3/c \\ E_1/c & 0 & -B_3 & B_2 \\ E_2/c & B_3 & 0 & -B_1 \\ E_3/c & -B_2 & B_1 & 0 \end{pmatrix}\,, \tag{1.10}$$

with $\mu, \nu = 0, \ldots, 3$. In this way, the two potential relationships (1.2), (1.3) can be replaced by:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \tag{1.11}$$

with the contra-variant gradient $(\partial^\mu) = (c^{-1}\partial_t, -\partial_1, -\partial_2, -\partial_3)$. The remaining two equations, (1.1c), (1.1d), can consequently be written in the elegant form $\partial_\mu F^{\mu\nu} = \mu_0 j^\nu$, where $(j^\nu) = \left(c\rho, \vec{j}\right)$ denotes the 4-vector of the current density.

Obviously, the above potential representation, together with proper gauging, is associated with several advantages. First of all, compared to (1.1a)–(1.1c), the set of equations (1.8a), (1.8b) is reduced, removing the redundancy associated with need of having four equations for four unknown fields. The latter form, comprising two wave equations, is

also more convenient in the sense that it is decoupled, self-adjoint and thus admits a least-action principle which was hitherto unavailable.

The representation of electromagnetism via potential quantities which are determined only up to a certain transformation (1.5), (1.6) left Maxwell's contemporaries uncomfortable initiating debates which were not free of polemics. Heaviside (1850-1925) seemed to have derived great satisfaction in 1885 in eliminating the vector potential $\vec{A}$ from the Maxwell equations. Disagreeing with Maxwell's elevation of $\vec{A}$ to the rank of a fundamental quantity, Heaviside ([121], Vol. 2, p. 482) regarded:

> "*$\vec{A}$ and its scalar potential* parasite *φ sometimes causing great mathematical complexity and indistinctiveness; and it is, for practical reasons, best to murder the whole lot, or, at any rate, merely employ them as subsidiary functions. [...] Thus φ and $\vec{A}$ are murdered, so to speak, with a great gain in definiteness and conciseness.*"

In a similar vein, Heaviside ([122], p. 383) had this to say:

> "*The reader who is acquainted with the (at present) more 'classical' method of treating the electromagnetic field in terms of the vector and scalar potentials cannot fail to be impressed by the difference of procedure and of ideas involved. In the present method we are, from first to last, in contact with those quantities which are believed to have physical significance (instead of with mathematical functions of an essentially indeterminate nature), and with the laws connecting them in simplest form.*"

Heaviside and Hertz correctly realised that Maxwell's theory could be completely described using the relevant observables, $\vec{E}$ and $\vec{B}$, and the potentials indeed allow for an elegant, possibly computationally useful, description which in their opinion, however, caused an undesirable "obscuring" [283] of the underlying physical laws. While their perception of the potentials spread in the scientific community, it was only with the advent of modern quantum field theory starting with Schrödinger's paper in 1926 that the deep significance of gauge invariance and correspondingly a different view on the electrodynamic potentials emerged, as outlined below.

In a non-relativistic quantum mechanics context consider the Schrödinger Lagrange density of a free particle in the *absence* of electromagnetic fields [141]:

$$\ell_0\left(\psi, \psi^*, \partial_\alpha\psi, \partial_\alpha\psi^*\right) = \frac{\hbar}{2\mathrm{i}}\left[\psi\partial_t\psi^* - \psi^*\partial_t\psi\right] - \frac{\hbar^2}{2m}\nabla\psi\cdot\nabla\psi^*\,, \tag{1.12}$$

where $\hbar = h/2\pi$, as usual, denotes the reduced Planck constant and $\psi(x_i, t)$ the complex-valued wave function of the particle at position x and time t; according to Born's interpretation [141] $\psi^*\psi$, with ψ^* referring to the adjoint of ψ, delivers the probability

density of the particle. When variations of (1.12) are taken with respect to ψ^* and ψ the time-dependent Schrödinger equation and its adjoint are obtained:

$$\mathrm{i}\hbar\partial_t\psi + \frac{\hbar^2}{2m}\Delta\psi = 0\,,$$
$$-\mathrm{i}\hbar\partial_t\psi^* + \frac{\hbar^2}{2m}\Delta\psi^* = 0\,.$$

Obviously, gauge symmetries of the form $\psi \to \psi' = \psi \mathrm{e}^{\mathrm{i}\chi}$ lead to the transformation behaviour:

$$\partial_\alpha\psi \to \partial_\alpha\left[\psi\mathrm{e}^{\mathrm{i}\chi}\right] = \left[\partial_\alpha\psi + \mathrm{i}\psi\partial_\alpha\chi\right]\mathrm{e}^{\mathrm{i}\chi}\,,$$

and thus turn out to be symmetries of the Lagrange density (1.12) if and only if $\chi = \text{const}$, i.e. if the phase gauge is *global*. Now, the introduction of further physical effects, for instance due to the presence of electromagnetic fields, is intimately connected with the requirement of *local* gauge symmetry according to the formalism of minimal coupling [141, 142], which is seen as follows. If invariance with respect to a local symmetry with arbitrary gauge field $\chi = \chi\left(x^\alpha\right)$ is required, an extension of the Lagrangian density (1.12) is unavoidable. One possibility is a formal substitution of the contra-variant gradient:

$$\partial_\alpha \longrightarrow \partial_\alpha - \mathrm{i}\frac{q}{c}A_\alpha\,,$$

involving an electric charge q, the speed of light c, and new so-called compensating fields A_α, cf. (1.10), leading to the transformation rule:

$$\left\{\partial_\alpha - \mathrm{i}\frac{q}{c}A_\alpha\right\}\psi \to \left\{\partial_\alpha - \mathrm{i}\frac{q}{c}A'_\alpha\right\}\left[\psi\mathrm{e}^{\mathrm{i}\chi}\right] = \left[\partial_\alpha\psi - \mathrm{i}\left(\frac{q}{c}A'_\alpha - \partial_\alpha\chi\right)\psi\right]\mathrm{e}^{\mathrm{i}\chi}\,.$$

This delivers the desired form $\left\{\partial_\alpha - \mathrm{i}\frac{q}{c}A_\alpha\right\}\psi \to \left\{\partial_\alpha - \mathrm{i}\frac{q}{c}A_\alpha\right\}\psi\mathrm{e}^{\mathrm{i}\chi}$ if and only if the compensating fields obey the transformation rule:

$$A_\alpha \to A'_\alpha = A_\alpha + \frac{c}{q}\partial_\alpha\chi\,,$$

which is exactly the, already known, gauge transformation for the electrodynamic Maxwell potentials (1.5), (1.6). The extended Lagrangian density can furthermore depend on invariants which can be formed from the compensating fields and their derivatives, particularly $\partial_\alpha A_\beta - \partial_\beta A_\alpha$, which in the simplest case (when squared) leads to the Lagrangian density $\ell_M\left(\partial_\alpha A_\beta - \partial_\beta A_\alpha\right)$ of Maxwell's theory, i.e. equation (1.4).

In summary the Lagrangian density:

$$\begin{aligned}\ell\left(\psi,\psi^*,A_\beta,\partial_\alpha\psi,\partial_\alpha\psi^*,\partial_\alpha A_\beta\right) = {}& \ell_0\left(\psi,\psi^*,\left\{\partial_\alpha - \mathrm{i}\frac{q}{c}A_\alpha\right\}\psi,\left\{\partial_\alpha + \mathrm{i}\frac{q}{c}A_\alpha\right\}\psi^*\right)\\ &+ \ell_M\left(\partial_\alpha A_\beta - \partial_\beta A_\alpha\right)\,,\end{aligned} \tag{1.13}$$

of Schrödinger-Maxwell theory for the motion of charged particles in electromagnetic fields is obtained which, by variation with respect to ψ, ψ^*, results in the following equation [142]:

$$\mathrm{i}\hbar\partial_t\psi + \hbar q\phi\psi + \frac{\hbar^2}{2m}\left\{\partial_k + \mathrm{i}\frac{q}{c}A_k\right\}^2\psi = 0\,, \tag{1.14}$$

and its adjoint, and by variation with respect to A_α in the Maxwell equations $\partial_\mu F^{\mu\nu} = \mu_0 j^\nu$ (as shown earlier). Interestingly, in the above formalism the Maxwell potentials occur naturally as "compensating fields" in order to allow for a local gauge symmetry and these compensating fields introduce the desired electromagnetic effects; here and also in the standard model of particle physics the, formerly objected to, gauge freedoms are elevated to a principle for the construction of theories. For further background to the development of gauge theory, Yang-Mills theory and the standard model, the reader is referred to [141].

Although with the above development physicists became aware of the conceptual significance of potential fields, in electrodynamics it remained an accepted standard that the electromagnetic field resides in the observables $\vec{E}$ and $\vec{B}$, i.e. where both of them vanish there cannot be any electromagnetic effects on a charged particle. This perception was attacked in 1959 in the article by Aharonov and Bohm [5] demonstrating an effect which, according to their explanation, is only comprehensible in the context of quantum electrodynamics when the vector potential is present. Their interpretation amounts to attributing a physically fundamental role to the potentials, which seems in contradiction to their gauge freedoms and implies that a representation of electrodynamics in terms of the potentials is somewhat physically more complete while, in the words of Wu and Yang [283], *"[...] the field strengths underdescribe electromagnetism"*. However, while the Aharanov-Bohm effect itself has been experimentally verified several times, cf. [283], the explanation remains a source of debates: typical contrary positions are described by Aharonov et al. [6], Wu and Yang [283] in favour of a quantum topological explanation and Boyer [42] preferring an explanation by classical electrodynamic effects.

While the Aharonov-Bohm effect is well described by quantum electrodynamics a frequent misleading argument attributes significance to the potentials because they cannot be eliminated from quantum electrodynamics in contrast to classical electrodynamics [283]; however, in 1927 Madelung [168, 201] demonstrated a gauge invariant hydrodynamic form of the equations which is reviewed briefly below as it leads to the focus of the present work: fluid mechanics.

By employing the substitution:

$$\varrho = m\psi^*\psi\,, \qquad \phi = \frac{\hbar}{m}\arg\psi\,, \qquad \psi = \sqrt{\frac{\varrho}{m}}\exp\left(\mathrm{i}\frac{m}{\hbar}\phi\right)\,,$$

in (1.13), it follows that:

$$\left\{\partial_\alpha - \mathrm{i}\frac{q}{c}A_\alpha\right\}\psi = \left\{\partial_\alpha - \mathrm{i}\frac{q}{c}A_\alpha\right\}\left[\sqrt{\frac{\varrho}{m}}\exp\left(\mathrm{i}\frac{m}{\hbar}\phi\right)\right]$$
$$= \left[\frac{\partial_\alpha\varrho}{2\varrho} + \mathrm{i}\frac{m}{\hbar}\partial_\alpha\phi - \mathrm{i}\frac{q}{c}A_\alpha\right]\psi\,,$$

giving an alternative form for the Lagrange density (1.13):

$$\ell = -\varrho\left[\frac{\partial\phi}{\partial t} - \frac{\hbar q}{m}A_0 + \frac{1}{2}\left(\nabla\phi - \frac{\hbar q}{mc}\vec{A}\right)^2\right] - \frac{\hbar^2(\nabla\varrho)^2}{8m^2\varrho} + \ell_M\left(\cdots\right)\,.$$

Variation of the corresponding action functional with respect to ϕ leads to a continuity equation of the form:

$$\frac{\partial\varrho}{\partial t} + \nabla\cdot\left(\varrho\left[\nabla\phi - \frac{\hbar q}{mc}\vec{A}\right]\right) = 0\,, \tag{1.15}$$

and therefore allows identification of the quantity in square brackets as a flow (velocity) field:

$$\vec{u} := \nabla\phi - \frac{\hbar q}{mc}\vec{A}\,, \tag{1.16}$$

while variation with respect to ϱ delivers a generalised Bernoulli equation:

$$-\left[\frac{\partial\phi}{\partial t} - \frac{\hbar q}{m}A_0 + \frac{1}{2}\left(\nabla\phi - \frac{\hbar q}{mc}\vec{A}\right)^2\right] + \underbrace{\frac{\hbar^2(\nabla\varrho)^2}{8m^2\varrho^2} + \nabla\cdot\left[\frac{\hbar^2\nabla\varrho}{4m^2\varrho}\right]}_{\frac{\hbar^2\Delta\sqrt{\varrho}}{2m^2\sqrt{\varrho}} =: -P} = 0\,.$$

The above can be written compactly as:

$$\frac{\partial\phi}{\partial t} - \frac{\hbar q}{m}A_0 + \frac{1}{2}\vec{u}^2 + P = 0\,, \tag{1.17}$$

and can be formally interpreted as a first integral of a "kind of Euler equation". This can be seen by forming the gradient of (1.17) upon using (1.16):

$$\frac{\partial}{\partial t}\left[\vec{u} + \frac{\hbar q}{mc}\vec{A}\right] - \frac{\hbar q}{m}\nabla A_0 + \frac{1}{2}\nabla\vec{u}^2 + \nabla P = 0\,,$$

which, via consideration of the vector identity $\frac{1}{2}\nabla\vec{u}^2 = (\vec{u}\cdot\nabla)\vec{u} + \vec{u}\times(\nabla\times\vec{u})$ together with the Maxwell relationships:

$$\nabla\times\vec{u} = -\frac{\hbar q}{mc}\nabla\times\vec{A} = -\frac{q}{m}\vec{B}\,, \tag{1.18}$$

$$-\frac{\hbar q}{m}\left[\nabla A_0 - c^{-1}\partial_t\vec{A}\right] = -\frac{q}{m}\vec{E}\,, \tag{1.19}$$

leads to:

$$\mathrm{D}_t\vec{u} = -\nabla P + \frac{q}{m}\left[\vec{E} + \vec{u}\times\vec{B}\right] . \tag{1.20}$$

Obviously, the Euler-like equations (1.20), (1.15) with the substitution (1.16) and complemented by the classical Maxwell equations (1.1a)-(1.1d) provide a representation purely in terms of the gauge invariant quantities $\vec{u}$, P, $\vec{E}$ and $\vec{B}$.

As a concluding remark one can state that, independent of the long-standing debate on the physical significance of potentials, a direct comparison between the gauge invariant equations (1.20), (1.15) and the original gauge variant Schrödinger equations (1.14) reveals beyond doubt a greater elegance and simplicity of the latter, not least due to its linearity and self-adjointness. This perception is also the origin of Anthony's work [9] employing a generalised Schrödinger representation of fluid mechanics, an idea which in contrast to the Madelung representation is far less familiar in the literature.

1.2 Potentials in fluid dynamics and the limits of Lagrange formalism

1.2.1 Potentials and first integrals of the equations of motion

Similar to electrodynamics[2] the use of auxiliary potential fields for different purposes has a long-standing history [99]. Consider, as a basis, the motion of a viscous, incompressible fluid governed by the Navier-Stokes and continuity equations [154]:

$$\varrho\left(\frac{\partial\vec{u}}{\partial t} + (\vec{u}\cdot\nabla)\,\vec{u}\right) = -\nabla p + \eta\nabla^2\vec{u} + \vec{f}\,, \tag{1.21a}$$

$$\nabla\cdot\vec{u} = 0\,, \tag{1.21b}$$

in which $\vec{u}$ denotes the velocity field, p the pressure and η, ρ and $\vec{f}$ the dynamic viscosity, density and external forces, respectively. Equations (1.21a) and (1.21b) represent momentum and mass balance in which the former is a version of Newton's second law with the inertial terms (per volume) stated on the left-hand side and the forces per volume, i.e. pressure, viscous and external forces, on the right-hand side. Two important special cases arise for flows which are either dominated by viscosity or by convection; in the former case the inertial terms can be neglected in favour of the viscous ones resulting in the *Stokes equations*, while in the latter case the viscous term $\eta\nabla^2\vec{u}$ is neglected resulting in the *Euler equations*, mentioned earlier.

Certainly equations (1.21a), (1.21b) differ significantly from Maxwell's equation (1.1a)-

[2] Consider also the work of Kambe [142, 143] for a comparison between electrodynamics and fluid dynamics and the role of gauge principles.

(1.1d), not only in formal issues such as the non-linearity involved but also in their theoretical foundation: the former equations are Galilean-invariant while the latter are Lorentz-invariant. Nevertheless the expectation is that potential fields in fluid mechanics can be utilised to a similar promising effect as in electrodynamics, both with respect to analytical and computational simplification, and to the development of a Hamiltonian formulation. However, the historical development has been tardy and disparate, mostly covering special cases; it is important to consider this in more detail in order to understand the context of this thesis.

While the governing equations of fluid mechanics in their modern form (1.21a), (1.21b) were derived by Euler (1752), Navier (1823) and Stokes (1845) [77] spanning a century, already Euler himself and most notably Lagrange (1781), cf. [140], introduced the notion of *potential flow* which is certainly the most familiar use historically of a potential field in fluid mechanics. According to the Helmholtz decomposition introduced in 1858 [123] any vector field can be resolved into the sum of an irrotational (curl-free) and a solenoidal (divergence-free) part, i.e. $\vec{u} = \nabla\phi + \nabla \times \vec{\psi}$, which when applied to the velocity field gives rise to Lagrange's potential formulation if $\vec{\psi} = 0$ and to the streamfunction formulation for incompressible flow if $\phi = 0$. Obviously both approaches allow for the reduction of unknown fields at the expense of a higher differential order leading to further analytical and numerical methods of solution; particularly the scalar streamfunction approach in two dimensions established as one of the most useful standard representations in this respect [154].

(a) Clebsch-like approach

It is illuminating to interpret the above potentials as tools to enable the *integration* of the equations of motion; the most familiar example is Bernoulli's equation which is obtained as a first integral of Euler's equations in the absence of vorticity and viscosity if the velocity vector $\vec{u} = \nabla\phi$ is perceived as the gradient of a scalar potential. Application of the velocity gradient representation to Euler's equation yields:

$$\vec{0} = \frac{\mathrm{D}\vec{u}}{\mathrm{D}t} + \nabla[P + U] = \nabla\left[\frac{\partial\phi}{\partial t} + \frac{\vec{u}^2}{2} + P + U\right], \tag{1.22}$$

with $P = \int \varrho^{-1}\mathrm{d}p$ denoting the so-called pressure function, U the potential energy of the external force, i.e. $\vec{f} = -\nabla U$ is assumed, and the operator $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \vec{u}\cdot\nabla$ denotes the *material time derivative*. In 1856 Clebsch [64, 154] proposed a non-standard potential representation for the velocity field for the case of inviscid and rotational flow:

$$\vec{u} = \nabla\phi + \alpha\nabla\beta\,, \tag{1.23}$$

in terms of three scalar potentials ϕ, α and β, the so-called Clebsch variables; the resulting set of equations together with the continuity equation turn out to be self-adjoint, permitting a variational formulation [154]. From a mathematical viewpoint, the potential representation (1.23) is a decomposition of the velocity field into a curl-free part $\nabla\phi$ and a helicity-free part $\alpha\nabla\beta$. This decomposition is not unique; by applying the gauge transformation:

$$\begin{aligned} \phi \longrightarrow \quad \phi' &= \phi + f(\alpha,\beta,t)\,, \\ \alpha \longrightarrow \quad \alpha' &= g(\alpha,\beta,t)\,, \\ \beta \longrightarrow \quad \beta' &= h(\alpha,\beta,t)\,, \end{aligned} \tag{1.24}$$

an equivalent set of Clebsch variables ϕ', α', β' is given if and only if the functions f, g, h fulfill the two PDEs [233]:

$$\frac{\partial f}{\partial \beta} + g\frac{\partial h}{\partial \beta} = \alpha\,, \tag{1.25}$$

$$\frac{\partial f}{\partial \alpha} + g\frac{\partial h}{\partial \alpha} = 0\,. \tag{1.26}$$

The benefit of the Clebsch transformation becomes apparent, again by application to Euler's equations for inviscid flows:

$$\vec{0} = \frac{\mathrm{D}\vec{u}}{\mathrm{D}t} + \nabla[P+U] = \nabla\left[\frac{\partial\phi}{\partial t} + \alpha\frac{\partial\beta}{\partial t} + \frac{\vec{u}^2}{2} + P + U\right] + \frac{\mathrm{D}\alpha}{\mathrm{D}t}\nabla\beta - \frac{\mathrm{D}\beta}{\mathrm{D}t}\nabla\alpha\,. \tag{1.27}$$

Being basically of the form:

$$\nabla[\cdots] + [\cdots]\,\nabla\alpha + [\cdots]\,\nabla\beta = \vec{0}\,, \tag{1.28}$$

this vector equation can be decomposed according to:

$$\frac{\partial\phi}{\partial t} + \alpha\frac{\partial\beta}{\partial t} + \frac{\vec{u}^2}{2} + P + U = F(\alpha,\beta,t)\,, \tag{1.29}$$

$$\frac{\mathrm{D}\alpha}{\mathrm{D}t} = -\frac{\partial F}{\partial\beta}\,, \tag{1.30}$$

$$\frac{\mathrm{D}\beta}{\mathrm{D}t} = \frac{\partial F}{\partial\alpha}\,, \tag{1.31}$$

with an unknown function $F(\alpha,\beta,t)$. By making use of the gauge transformation (1.24), $F \to 0$ is achieved and (1.29) takes the form of a generalised Bernoulli equation. The above three scalar field equations are a first integral of Euler's equations and self-adjoint, their most intriguing feature is that the vorticity:

$$\vec{\omega} = \frac{1}{2}\nabla\times\vec{u} = \frac{1}{2}\nabla\alpha\times\nabla\beta\,, \tag{1.32}$$

is given by the two scalar fields α, β, only. Hence, the vortex dynamics is reduced to two transport equations (1.30), (1.31).

It should be noted that for an arbitrary velocity field $\vec{u}$ the existence of the Clebsch variables ϕ, α, β is guaranteed only locally. Their global existence depends on the topological features of the flow: in the case of a non-vanishing integral of helicity, for example for flows with closed vortex lines that form linked rings or with isolated points of zero vorticity, global existence is not given. For further details the reader is referred to [20, 286]. In the case of global non-existence, completeness of the Clebsch representation may be reached by additional pairs of variables, like: $\vec{u} = \nabla\phi + \alpha_1\nabla\beta_1 + \alpha_2\nabla\beta_2$. The Clebsch transformation has also been applied to different physical problems, for instance to baroclinic flow [236], Maxwell's equations in classical electrodynamics [271], to Magnetohydrodynamics [58] and even quantum theory within the context of a quantization of vortex tubes [168]. Viscous flow in terms of Clebsch variables is considered in this thesis for the first time.

(b) Goursat-like approach

A range of potential field methods for viscous flow based on the biharmonic streamfunction representation of the Stokes equations has arisen from the fruitful exchange of ideas between linear elasticity and fluid mechanics. These methods aim at expressing physical quantities in terms of more tractable functions, examples being the Papkovich-Neuber representation of velocity or the complex Goursat representation of the streamfunction which only depend on harmonic and analytic functions, respectively [111, 154]; as a consequence this reduction allows the employment of the whole of holomorphic function theory and their solution tools. However, there is also a deep relation to the above idea of integrating the equations of motion, as will become clear below.

In the first half of the 20th century the complex variable method was developed for the solution of biharmonic problems in plane linear elasticity, which nowadays is considered a classical [177, 184] approach. By combining the two Cartesian coordinates x and y to form the complex expression $\xi = x + \mathrm{i}y$, all solutions of the biharmonic equation $\Delta^2\Psi = 0$ take the form:

$$\Psi = \mathrm{Re}\left[g_0(\xi) + \bar{\xi}g_1(\xi)\right] , \tag{1.33}$$

with arbitrary holomorphic functions g_0 and g_1, the so-called *Goursat functions*; Re denotes the real part of the subsequent complex expression and $\bar{\xi} = x - \mathrm{i}y$ the complex conjugate of ξ. The problem is thus effectively reduced to one of finding two holomorphic functions in a domain Ω, with boundary $\partial\Omega$, satisfying conditions which are typically of the form [184]:

$$g_1(\xi) + \xi\overline{g_1'(\xi)} + \overline{g_0'(\xi)} = f(\xi) , \quad \xi \in \partial\Omega . \tag{1.34}$$

One such approach is to represent g_0 and g_1 as Cauchy-type integrals of an unknown density function on the boundary, thus leading to the well-known Sherman-Lauricella equation [84], a Fredholm integral equation for the density which can be solved efficiently via a Nyström discretization in combination with fast multipole methods [111]. Other approaches are based on a direct expansion of the holomorphic functions as a Fourier series or in meromorphic functions with singularities outside of the domain, in which (1.34) is approximated via boundary element or collocation methods [70, 144, 194, 217]. Although the biharmonic equation is not preserved under conformal transformation, diverse examples exist where a mapping of the Goursat representation to computationally more convenient regions has proved useful [60, 184].

Beyond linear elasticity, complex variable methods have emerged as a wide-ranging and powerful tool in fluid mechanics [187, 194], where a variety of analytical solutions and computational methods have materialised, particularly for Stokes flow problems with free boundaries [131, 202, 227, 228, 247]. The benefits thereof stem from the mathematically important notion of the analyticity of the underlying holomorphic functions g_0 and g_1 which can be exploited for certain problems [70, 194]; free surface boundary conditions, a notorious complicating feature, can be handled in the framework of the Goursat representation naturally, due in the main to the complementary character of the streamfunction and Airy stressfunction[3] [8, 171, 190, 251]. Unfortunately, the method in its classical form is restricted to zero Reynolds number scenarios, i.e. to Stokes flow, in which the streamfunction remains biharmonic.

More recently, an approach to integrate the steady 2D-NS equations has been established by Ranger [199] and Scholle et al. [230]. There it is shown that a complex variable transformation $\xi := x+\mathrm{i}y$, $\overline{\xi} := x-\mathrm{i}y$ of the field equations, together with the introduction of a complex velocity field $u := u_x + \mathrm{i}u_y$, yields the equivalent formulation[4]:

$$\frac{\partial}{\partial\overline{\xi}}\left[p + \varrho\frac{\bar{u}u}{2} + U\right] + \varrho\frac{\partial}{\partial\xi}\left(\frac{u^2}{2}\right) = 2\eta\frac{\partial^2 u}{\partial\overline{\xi}\partial\xi} , \tag{1.35a}$$

$$\mathrm{Re}\left(\frac{\partial u}{\partial\xi}\right) = 0 , \tag{1.35b}$$

where Re denotes the real part of the subsequent complex expression. The introduction of a scalar real-valued potential $\Phi(\xi,\overline{\xi})$ satisfying:

$$p + \varrho\frac{\bar{u}u}{2} + U = 4\frac{\partial^2\Phi}{\partial\overline{\xi}\partial\xi} , \tag{1.36}$$

[3] This is explained in more detail in Sec. 3.1 of Chapter 3.

[4] Note, that in contrast to [230] the standard complex variable transformation is used here leading to slightly different equations.

allows (1.35a), in the first place, to be recast in an integrable fashion:

$$\frac{\partial}{\partial\xi}\left[\varrho\frac{u^2}{4}-\eta\frac{\partial u}{\partial\bar{\xi}}+2\frac{\partial^2\Phi}{\partial\bar{\xi}^2}\right]=0\ , \tag{1.37}$$

and in the second place to be integrated with respect to ξ, delivering a result dependent on an undetermined integration function $g(\bar{\xi})$:

$$\varrho\frac{u^2}{4}-\eta\frac{\partial u}{\partial\bar{\xi}}+2\frac{\partial^2\Phi}{\partial\bar{\xi}^2}=g(\bar{\xi})\ . \tag{1.38}$$

At this point the close relationship between the terms on the left-hand side of equation (1.36) and the steady version of Bernoulli's equation is observed, indicating that those terms will vanish in the case of irrotational and inviscid flow. In the general case of a viscous flow with non-vanishing vorticity, (1.36) can be interpreted as a potential representation of the deviation from Bernoulli's equation.

Further simplification follows by observing that the potential field Φ itself is not uniquely determined by (1.36) but only up to the sum of an arbitrary holomorphic function $f(\xi)$ and its complex conjugate:

$$\Phi\rightarrow\Phi+f(\xi)+\overline{f(\xi)}\,, \tag{1.39}$$

which facilitates the elimination of g by fixing the gauge freedom (1.39) according to $f''(\bar{\xi})=g(\bar{\xi})$. Finally an alternative to the classical representation (1.21a), (1.21b) of the NS equations is given by the integrated set of equations:

$$\varrho\frac{u^2}{4}-\eta\frac{\partial u}{\partial\bar{\xi}}+2\frac{\partial^2\Phi}{\partial\bar{\xi}^2}=0\ , \tag{1.40a}$$

$$\mathrm{Re}\left(\frac{\partial u}{\partial\xi}\right)=0\ . \tag{1.40b}$$

Note, the resulting equations contain first order derivatives only of velocity in contrast to the original NS equation (1.21a)-(1.21b) and the corresponding 2D streamfunction version in general contains second order derivatives only, whereas the classical streamfunction formulation results in a fourth order equation[5].

A real-valued representation of the complex system (1.40a)-(1.40b) in tensor notation reads:

$$\eta\left[\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}-\frac{\partial u_k}{\partial x_k}\delta_{ij}\right]-\varrho\left[u_iu_j-u_ku_k\frac{\delta_{ij}}{2}\right]=2\left[\frac{\partial^2\Phi}{\partial x_i\partial x_j}-\frac{\partial^2\Phi}{\partial x_k\partial x_k}\frac{\delta_{ij}}{2}\right], \tag{1.41}$$

[5]Here the case applies that is mentioned at the beginning of the introduction: eq. (1.37) facilitates a first integral of (1.35a) while, due to the introduction of Φ, the total differential order in (1.40a) remains two.

together with:

$$\frac{\partial u_k}{\partial x_k} = 0 \ , \qquad p + \frac{\varrho}{2} u_k u_k + U = \frac{\partial^2 \Phi}{\partial x_k \partial x_k} \ . \tag{1.42}$$

The second equation in (1.42) is only relevant in applications where the recovery of pressure is of interest; the pressure can easily be computed subsequently.

In a recent paper Marner et al. [171] revealed the close relationship between the first integral representation and the complex variable method: if a streamfunction Ψ is introduced according to:

$$u = -2\mathrm{i}\frac{\partial \Psi}{\partial \bar{\xi}} \ , \tag{1.43}$$

it can be beneficially combined with the potential Φ to form the complex potential, $\chi = \Phi + \mathrm{i}\eta\Psi$, in terms of which the first integral (1.40a) reads:

$$\frac{\partial^2 \chi}{\partial \bar{\xi}^2} = -\varrho \frac{u^2}{8} \ . \tag{1.44}$$

In the case of Stokes flow equation (1.44) reduces to the elegant form:

$$\frac{\partial^2 \chi}{\partial \bar{\xi}^2} = 0 \ , \tag{1.45}$$

which is a simple bianalytic equation; the solution of which is given by the combination $\chi = g_0(\xi) + \bar{\xi} g_1(\xi)$ of two holomorphic functions, as in (1.33).

A key feature of the above is it can be construed that the first integral (1.44) is a generalisation of the complex-valued potential formulation (1.45) towards viscous flows with inertia, with the difference that (1.44) no longer allows for a direct integration to the Goursat representation (1.33). Based on this first integral of the equations of motion, a non-conventional variational formulation has been established in order to solve for plane creeping flow by a semi-analytical direct Ritz method, Scholle et al. [230]; further contributions are a contour integral method developed by Coleman [70] and a least-squares approach by Bolton and Thatcher [38]. Most notably, a similar set of equations was used by Ranger [199] to achieve a parametrized general solution of the NS equations; he later demonstrated a generalisation to three-dimensional (3D) flow [198], additionally see his other papers and references therein.

Similar to electrodynamics the representation of physical quantities in terms of potential fields is closely related to the construction of variational principles. This is best known for ideal fluids as shown below.

1.2.2 Potentials and variational principles

The goal of describing processes in terms of maximum and minimum principles is an alluring one which has occupied scientists for centuries. The first such principles were Fermat's principle of least time in optics and Hamilton's principle of least action in mechanics [213]. In 1744 the mathematician Euler expressed the hope, cf. [99]:

> "*As the construction of the universe is the most perfect possible, being the handiwork of an all-wise Maker, nothing can be met with in the world in which some maximal or minimal property is not displayed. There is, consequently, no doubt but that all the effects of the world can be derived by the method of maxima and minima from their final causes as well as from their efficient ones.*"

A concise mathematical manifestation of the above statement is provided by the Lagrange formalism [216] of field theory which also comprises the least action principle of classical mechanics as a special case. Consider a dynamical physical system described by a set of fields $\psi = (\psi_1, \ldots, \psi_n)$ depending on generalised coordinates q_k and time t, i.e. $\psi_i = \psi_i(q_k, t)$, then *Hamilton's principle* for fields states that among all possible virtual processes in a fixed domain Ω with fixed endpoints $\psi_i(q, t_0) = g_0(q_k)$, $\psi_i(q, t_1) = g_1(q_k)$ the effective process taking place in the time interval $[t_0, t_1]$ is distinguished by enforcing the stationarity (not necessarily the extremum) of a corresponding *action functional*:

$$L[\psi_i] = \int_{t_0}^{t_1} \int_{\Omega} \ell(\psi_i, \partial_k \psi_i, \partial_t \psi, q_k, t) \, \mathrm{d}\Omega \, \mathrm{d}t \,. \tag{1.46}$$

Here, the functional L is also called the *Lagrangian* and the corresponding integrand ℓ the *Lagrangian density*. Setting the functional derivatives to zero, i.e.:

$$\frac{\delta L[\psi]}{\delta \psi_i} = 0 \,,$$

implies for $i = 1, \ldots, n$ the *Euler-Lagrange equations* [108]:

$$\frac{\partial \ell}{\partial \psi_i} - \frac{\partial}{\partial t}\left(\frac{\partial \ell}{\partial \dot{\psi}_i}\right) - \nabla \cdot \left(\frac{\partial \ell}{\partial \nabla \psi_i}\right) = 0 \,. \tag{1.47}$$

The classical Hamilton's principle of particle mechanics results if the fields ψ_i are simply replaced by the Cartesian coordinates x_i of n particles. For these particles, usually the kinetic and potential energy are obtained as $T = \frac{1}{2}\sum_i m_i \dot{x}_i^2$ and $U = U(x_i)$ and the action functional (1.46) is replaced by:

$$L[x_i] = \int_{t_0}^{t_1} T(\dot{x}) - U(x) \, \mathrm{d}t \,, \tag{1.48}$$

leading to Newton's second law for a conservative force as Euler-Lagrange equations:

$$m_i \ddot{x}_i = -\frac{\partial U}{\partial x_i}\,, \quad i = 1, \ldots, n\,.$$

Since Euler's famous statement and the advent of Hamilton's principle, variational laws in nature have attracted the attention of many scientists and this attraction is certainly not only from a philosophical or aesthetic point of view. It is insightful to distinguish how the interest in such principles varies from the standpoint of different scientific communities and how this determines the assessment of the usefulness of them, see [254]. First of all, the main interest of the applied mathematician in the variational formulation lies in the fact that it offers a method of *proving information* about the existence of the solution of a given problem; here one speaks of "energy arguments", not necessarily referring to the physical energy [92]. Similar methods are also useful when establishing stability bounds in the field of hydrodynamics, as in the case of two superposed fluids with different densities [61] or water waves [167]. Contrastingly, the main interest of the engineer in variational formulations lies in the possibility of utilising numerical methods of solution and thus his/her aim is the *finding* of the solution. The so-called direct Ritz method of the calculus of variations, both in a semi-analytic or finite element like fashion, generally leads to favourably structured linear systems allowing for efficient solution strategies; the relationship between variational principles and Galerkin methods is well described by Finlayson and Scriven [97] in the unifying framework of the method of weighted residuals. Moreover, in the case of Hamiltonian systems symplectic integration schemes become available which numerically preserve physical conservation quantities [161] and are particularly suitable for long-term simulations of evolution equations; more than standard Runge-Kutta schemes. The main interest of the physicist is different still. S/he is fascinated by the synthetic statement and by the physical meaning of the functional (it may be a time, an energy, an action, a transition probability, etc.). But what seems to to be the crucial impetus of the physicist, is the axiom that equations obtained from a variational principle constitute a privileged class among all conceivable equations. This is probably due to the fact that many of the field equations are derived from a variational principle [254] and, more importantly, that the Lagrange formalism elegantly allows the establishment of an immediate connection between symmetry principles and physical conservation laws via the unified framework of Noether's theorem [155, 218, 235].

The wide and manifold interest in variational principles has initiated the so-called "inverse problem of the calculus of variations" [150, 214, 218] which is concerned with the question whether a system of differential equations is variational, i.e., if it can be expressed in the form of Euler–Lagrange equations for some Lagrangian, and how this Lagrangian can be found. The original (restricted) form of the problem given for instance by Hirsch

[128] and cited from Tonti [254] is:

Definition 1.1. *Given a (non-linear) problem $N(u) = 0$ with $N : D(N) \subset U \to R(N) \subset V = U^*$, find a functional $L[u]$, if any, such that the operator N is the gradient of L, i.e. such that:*

$$L'_u\varphi := \lim_{\varepsilon\to 0} \frac{L(u+\varepsilon\varphi) - L(u)}{\varepsilon} = (N(u), \varphi)\,.$$

This implies that the solutions of problem $N(u) = 0$ are the critical points of the functional $L[u]$ and vice versa.

In order that an operator N be the gradient of a functional L it is necessary that the circulation of the element $v = N(u)$ along any reducible closed line contained in $D(N)$ vanishes which requires the equations $N(u) = 0$ to be self-adjoint or the operator N'_u to be symmetric, respectively, expressed by:

$$(N'_u\varphi, \psi) = (N'_u\psi, \varphi)\,. \tag{1.49}$$

The necessary symmetry condition (1.49) becomes sufficient if the domain $D(N)$ is simply connected, a classical result by Volterra [270] and rigorously proved by Vainberg [261]. An alternative form of (1.49) in terms of differential conditions dating, at least, back to Helmholtz, cf. [150], has been widely used to test differential equations for self-adjointness as for instance demonstrated by Finlayson [98].

However, there are physical systems, mostly of a dissipative nature, that have remained resistant for a long time to a variational formulation and apparently do not admit a Lagrangian according to Def. 1.1; this is the case for the NS equations of fluid dynamics, Fourier's equations of heat conduction and others. Though still optimistic Helmholtz restricted Euler's original statement when commenting in his classical paper on the principle of minimal action [125]:

> "*Already now one can consider highly probable that such a principle of minimal action represents the general law of all reversible processes in nature.*"

But irrespective of the persistent belief that the Lagrange formalism is *confined* to the above mentioned reversible processes, propagated for example by a theorem of Bauer [26], already Bateman [23, 24] presented a counter example in 1929 demonstrating that in some cases dissipative systems are amenable also, at least when altering them in a suitable way. The existence of equations which well describe physical phenomena although they are not deducible from a variational principle have stimulated physicists and engineers to *extend* the classical calculations of variations. In this vein Davis [80], in 1928, formulated an extended version of the inverse problem, namely to find a functional whose critical points

are solutions of the problem and vice versa; this statement is more general compared to the former because it requires only the coincidence of the critical points with the solution, without the supplementary requirement that N be the gradient of the functional L.

The extension of the classical variational calculus has led to a confusing variety of approaches which is difficult to classify; however, such a classification would be desirable in order to position a new potential-based variational formulation of the Navier-Stokes equations as provided in this thesis. Certainly one of the oldest methods is that of *Lagrange multipliers*: if a Lagrangian is known which reproduces parts of the desired system of equations a typical approach is to add further equations via so-called Lagrange multipliers [214]; these multipliers appear as new unknown-fields which by variation produce the desired equations. However, this extends the number of fields and variation with respect to the original physical fields delivers further adjoint equations which in total amounts to adding the complete adjoint problem; then one talks of the adjoint method [99] offering quite generally stationary principles for non-linear equations. So why is this not the solution of the inverse problem of variational calculus? Certainly such methods have their eligibility in optimisation theory [244], but normally the engineer is concerned by the increased number of unknown fields and equations preferring any other kind of Galerkin method while, certainly, the physicist is unsatisfied with the lack of physical meaning of the multipliers and, even more, with the frequently occurring non-causality of adjoint equations. At least sometimes a few of the multipliers can be eliminated or a physical meaning can be attached to these otherwise artificial fields as is the case for ideal flow, see Subsec. 1.2.2(c). In this vein Salmon [213] came to the conclusion:

> "*While the beauty of Hamiltonian theory may reside in its formal mathematical structure, its real importance depends very much on the physical meaning attached to the mathematical symbols.*"

Instead of adding further equations, equivalently the corresponding least-squares problem may be solved which is of higher differential order; this amounts to minimising the (physically artificial) residual in some suitable Hilbert norm and is in principle also applicable to general non-linear systems. Although the latter approach might be attractive from a numerical point of view and many of the essential advantages of a variational principle are present, the higher order in the variational formulation has confined it mostly to first order systems or first order reformulations allowing for the employment of standard continuous and piece-wise smooth finite elements [36].

An outstanding and one of the most extensive and general approaches to the inverse problem has been derived by functional analytic tools. Here, a fundamental observation is that the necessary symmetry condition (1.49) involves the use of a bilinear form and that symmetry of an operator as well as conservativeness of a vector field are properties *with*

respect to a bilinear form. This implies that if an operator does not satisfy the symmetry criterion, one may change the canonical bilinear form in order to ensure that the given non-linear operator does. Gurtin [116], Magri [169] and Tonti [253] show that for general linear initial value problems the employment of a convolution bilinear form allows for the construction of infinitely many variational formulations giving the desired equations in *equivalent* integro-differential form. Later Tonti [255] demonstrated that this approach is closely related to the method of integrating factors, known for its application to the damped harmonic oscillator [255], and later generalised it in the method of integrating operators [254, 255] which continues to be explored today [87]: actually a change of bilinear form can always be associated with an operator multiplication of the original equations. In [254] it is shown that, given a non-linear equation $N(u) = 0$ under certain conditions a suitable linear, symmetric and invertible integral operator K can be found such that the critical points of the functional $L[u] := \frac{1}{2}(N(u), KN(u))_0$, the gradient of which gives $\bar{N}(u) = N'^*_u KN(u)$, coincide with the solutions of $Nu = 0$[6].

Other approaches to the inverse problem of variational calculus include the generalised bracket formalism [113, 145, 183], the complex-valued field approach mainly promoted by Anthony [9, 10, 11][7], fractional derivatives [76, 204] and others. From a more physical point of view Scholle [216, 218] investigates the role of symmetries for the construction of Lagrangians. Also "quasi" variational formulations and "restricted" formulations have been suggested which frequently loose most or all of the advantages mentioned at the outset, cf. [99]. While a thoroughly researched historical treatise on the inverse problem of variation containing all the relevant older literature is presented by Kotulek [150] two nice reviews by Santilli [214], and from a differential geometric point of view Morandi et al. [182] cover the more modern development. Excellent sources for the development of variational principles in fluid mechanics are the book of Finlayson [99] and, particularly for the Hamiltonian description of ideal flow, Salmon [213] and well summarised in a more modern description by Morrison et al. [183].

The following focuses on the field of fluid mechanics with emphasis on variational principles which appear physically meaningful, following a path pursued by Anthony [9] and particularly Scholle [216, 218]. Finlayson [99] collects problems in fluid mechanics and heat transfer for which variational principles exist; in special cases the variational integral is a minimum or maximum, often the variational integral is only stationary, and in the general case no variational principle exists at all, except after inclusion of the adjoint problem or modifications according to other rules as outlined above.

[6] One has to be careful when replacing a problem $Nu = f$ by $N'^*_u Nu = N'^*_u f$ because generally the domain of the operator $N'^*_u Nu$ is a restriction of the domain of N; then the restriction of the domain necessary for the application of the operator N'^*_u would exclude elements f, that is possible source distributions. Thus the role of K is to modify the range of N making it "digestible" to the operator N'^*_u [255].

[7] Cf. the correspondence between the Schrödinger and Euler equations in Sec. 1.1.

(a) Simple principles in fluid mechanics

Initially, consider some simple examples of fluid dynamics which indeed allow for extremal principles and for a physically meaningful Lagrangian. Take (1.21a) and (1.21b) in their steady form with viscosity neglected, then among all incompressible velocity fields $\vec{u}$ in the domain Ω with, say, boundary conditions $\varrho\vec{u}\cdot\vec{n} = g$ on $\partial\Omega$ the irrotational motion of the form $\vec{u} = \nabla\phi$ has the least kinetic energy:

$$L_1[\vec{u}] = \frac{1}{2}\varrho\int_\Omega \vec{u}^2\,\mathrm{d}\Omega\,;$$

the so-called *Kelvin's principle*. On the other hand, among all irrotational fields $\vec{u} = \nabla\phi$ an incompressible flow with $\varrho\vec{u}\cdot\vec{n} = g$ maximises the functional:

$$L_2[\phi] = -\frac{1}{2}\varrho\int_\Omega \nabla\phi\cdot\nabla\phi\,\mathrm{d}\Omega + \int_{\partial\Omega}\phi g\,\mathrm{d}\partial\Omega\,.$$

Together there exists a reciprocal relationship $L_2(\phi) \leq$ kin. energy $\leq L_1[\vec{u}]$. Although these principles are "useful" they are related to rather specific cases and rarely applicable. At least there exist extensions to steady compressible flow according to Bateman [23] with the variational integral being the kinetic energy plus one-half the pressure, but they are not examined in detail here.

If now the viscous terms are present but the inertial terms are absent, equations (1.21a), (1.21b) reduce to the Stokes equations, say with velocity boundary conditions:

$$\eta\Delta\vec{u} - \nabla p = f \quad \text{in } \Omega\,, \tag{1.50a}$$

$$\nabla\cdot\vec{u} = 0 \quad \text{in } \Omega\,, \tag{1.50b}$$

$$\vec{u} = \vec{g} \quad \text{auf } \partial\Omega\,, \tag{1.50c}$$

and the *Helmholtz-Korteweg principle* from 1868 [124] applies, stating that among all incompressible flows $\vec{u}$ with boundary conditions (1.50c) the one fulfilling the Stokes equations (1.50a) has the least viscous dissipation. As an alternative to incorporating the incompressibility constraint via a restriction of the admissible space of vector fields, it can be appended to the functional via the method of Lagrange multipliers. This leads to the following Lagrangian density:

$$\ell(u_i, \partial_j u_i, p) = 2\eta D_{ij}D_{ij} - p\frac{\partial u_i}{\partial x_i}\,, \tag{1.51}$$

involving the strain rate tensor:

$$D_{ij} = \frac{1}{2}\left[\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right]\,.$$

In (1.51), the pressure p acts as a Lagrange multiplier for the incompressibility constraint (1.50b), i.e. (1.50b) results from variation with respect to p and (1.50a) from variation with respect to $\vec{u}$. However, as a drawback the original minimisation principle is converted to a stationary principle requiring much more advanced (saddle point) methods [28] to solve the corresponding linear systems. Variational principles for Stokes flow in terms of other variables, i.e. in terms of the streamfunction only or the streamfunction and vorticity, are analysed in detail by De Coster et al. [81] based on an investigation of the biharmonic problem.

There exist also a variety of variational principles for general unsteady ideal flow which is an essential ingredient for the subsequent treatment of the full Navier-Stokes equations and for this reason considered in the following paragraphs in more detail. Further generalisations to magnetohydrodynamics and non-Newtonian fluids are summarised by Finlayson [99], which are of no further relevance in the present context, while an analysis of general thermodynamic principles like the least entropy concept of Onsager [189] or related concepts [196, 208] are not universally valid. Although Finlayson [99] and others come to the conclusion that none of these physically meaningful principles are capable of incorporating both inertial and viscous effects at the same time, there remains the option to modify the theory of fluid mechanics itself to make it fit into the framework of Lagrange formalism. A relatively close form of modification is considered in the following, subsequent to an analysis of the ideal flow case.

(b) Fluid flow from two view points

Often, fluid mechanics is treated entirely in terms of Eulerian variables as in the description (1.21a), (1.21b), but the simplest form of Hamilton's principle for a perfect fluid is obtained by a straightforward generalization of (1.48) to the case of infinitely many marked particles distributed continuously in space [183, 213]. Thus before proceeding to discuss variational principles it is useful to clarify these two different views on fluid mechnics.

In the Lagrangian variable description, also called *material description*, one picks out a particular particle, labelled by $\vec{a}$, and keeps track of where it goes in time t. Suppose the position of a fluid element with a specified label $\vec{a}$ at a time t, referred to a fixed rectangular coordinate system, is given by Lagrangian coordinates $\vec{q} = (q_1, q_2, q_3)$, $\vec{q} = q(\vec{a}, t)$. The label $\vec{a} = (a_1, a_2, a_3)$ should uniquely identify the fluid element and is often taken to be the element's position at time $t = 0$, but obviously exhibits a gauge freedom. Assume that $\vec{a}$ varies over a fixed domain $\Omega \subset \mathbb{R}^3$ which is completely filled with fluid, and that for a fixed t the function $q\colon \Omega \to \Omega$ is one-to-one and onto. Moreover, assume that as many derivatives of q with respect to a as needed exist. Often in the Hamiltonian context the functions $\vec{q} = q(\vec{a}, t)$ are assumed to be diffeomorphisms and their collection is referred

to as the diffeomorphism group [183]. Viewing the map $\vec{a} \mapsto \vec{q}$ at fixed t as a coordinate change, the Jacobian matrix $J_{ij} = \partial q_i / \partial a_j$ has an inverse given by:

$$\frac{\partial q_k}{\partial a_j} \frac{A_{ki}}{\det J} = \delta_{ij} \,,$$

where A is the *cofactor matrix* of J, being conveniently defined through:

$$A_{ki} = \frac{1}{2} \varepsilon_{kjl} \varepsilon_{imn} \frac{\partial q_j}{\partial a_m} \frac{\partial q_l}{\partial a_n} \,. \tag{1.52}$$

Alternatively, in the so-called *Eulerian field description*, one stays at a spatial observation point $\vec{x} = (x_1, x_2, x_3) \in \Omega$ and monitors the nature of the fluid at $\vec{x}$ at time t. The most important Eulerian variable is the Eulerian velocity field $\vec{u}(\vec{x}, t)$. This quantity is the velocity of the particular fluid element that is located at the spatial point $\vec{x}$ at time t. The label of that particular fluid element is given by $\vec{a} = q^{-1}(\vec{x}, t)$, and so:

$$\vec{u}(\vec{x}, t) = \dot{q}(\vec{a}, t)\big|_{\vec{a} = q^{-1}(\vec{x}, t)} := \dot{q} \circ q^{-1}(\vec{x}, t) \,, \tag{1.53}$$

where $\dot{q}$ denotes differentiation with respect to time at fixed label $\vec{a}$. Attached to a fluid element is a certain amount of mass described by a density function $\varrho_0(\vec{a})$. As the fluid moves so that $\vec{a} \mapsto \vec{q}$, the volume of an infinitesimal region will change, but its mass must remain fixed. The statement of local mass conservation is $\varrho \mathrm{d}^3 x = \varrho_0 \mathrm{d}^3 a$, where $\mathrm{d}^3 a$ is an initial infinitesimal volume element that maps to $\mathrm{d}^3 q$ at time t, and $\mathrm{d}^3 x = \det(J) \mathrm{d}^3 a$[8], giving:

$$\varrho(\vec{x}, t) = \left. \frac{\varrho_0(\vec{a})}{\det J(\vec{a}, t)} \right|_{\vec{a} = q^{-1}(\vec{x}, t)} = \frac{\varrho_0}{\det J} \circ q^{-1}(\vec{x}, t) \,. \tag{1.54}$$

Besides the density, for an ideal fluid, one attaches an entropy per unit mass, $s = s_0(\vec{a})$, to a fluid element, and this quantity remains fixed in time. In the Eulerian description this gives rise to the entropy field:

$$s(\vec{x}, t) = s_0(\vec{a})\Big|_{\vec{a} = q^{-1}(\vec{x}, t)} = s_0 \circ q^{-1}(\vec{x}, t) \,. \tag{1.55}$$

Equations (1.53)-(1.55) express the Euler–Lagrange map.

(c) Ideal flow action principle in Lagrangian coordinates

Construction of the Lagrangian requires identification of the potential energy, and this requires thermodynamics, because potential energy is stored in terms of pressure and temperature. A basic assumption of the fluid approximation is that of local thermody-

[8] When integrating over Ω, $\mathrm{d}^3 q$ is replaced by $d^3 x$.

namic equilibrium. In the energy representation of thermodynamics, the extensive energy is treated as a function of the entropy and the volume. For a fluid, it is convenient to consider the energy per unit mass, denoted by e, to be a function of s and the mass density ϱ a measure of the volume. The intensive quantities, temperature and pressure, are given by:

$$T = \frac{\partial e}{\partial s} \quad \text{and} \quad p = \varrho^2 \frac{\partial e}{\partial \varrho} \,. \tag{1.56}$$

Special choices for e produce specific fluid flows like barotropic or adiabatic flow. Conventional thermodynamic variables can be viewed as Eulerian variables with a static velocity field. Writing $e(\varrho, s)$, where ϱ and s are spatially independent or, if the system has only locally relaxed, these variables can be functions of $\vec{x}$. For an ideal fluid, each fluid element can be viewed as a self-contained isentropic thermodynamic system that moves with the fluid. The total fluid potential energy functional and the kinetic energy functional are given by:

$$U(q) = \int_\Omega \varrho_0 e\left(\frac{\varrho_0}{\det J}, s_0\right) \mathrm{d}^3 a \quad \text{and} \quad T(q, \dot{q}) = \int_\Omega \varrho_0 \frac{\dot{q}^2}{2} \, \mathrm{d}^3 a \,,$$

respectively, in which the former is a functional of q that depends only upon $\det J$ and hence only upon $\partial q_i / \partial a_j$. The action functional analogous to (1.48) is then:

$$L = \int_{t_0}^{t_1} \int_\Omega \left[\frac{1}{2}\varrho_0 \dot{q}^2 - \varrho_0 e\left(\frac{\varrho_0}{\det J}, s_0\right)\right] \mathrm{d}^3 a \, \mathrm{d}t \,. \tag{1.57}$$

The end conditions for Hamilton's principle for the fluid are the same as those of mechanics, that is, $\delta q(a, t_0) = \delta q(a, t_1) = 0$. The non-penetration condition, $\delta q \cdot n = 0$ on $\delta\Omega$, where n is a unit normal vector is also assumed. Other boundary conditions, such as periodic and free boundary conditions, are also possibilities. Hamilton's principle amounts to $\delta L / \delta q(a, t) = 0$, which, together with the end and boundary conditions, implies the following equations of motion:

$$\varrho_0 \ddot{q}_i + A_{ij} \frac{\partial}{\partial a_j} \left(\frac{\varrho_0^2}{(\det J)^2} \frac{\partial e}{\partial \varrho}\right) = 0 \,. \tag{1.58}$$

Here, the identity $\partial A_{ij} / \partial a_j = 0$ has been used, which can be seen using (1.52). Equation (1.58) amounts to Newton's second law for an ideal fluid, which is made clearer by using the following useful identity:

$$\frac{\partial}{\partial q_i} = \frac{1}{\det J} A_{ij} \frac{\partial}{\partial a_j} \,. \tag{1.59}$$

The Eulerian variable force law follows from (1.58) upon using (1.59):

$$\varrho\left(\frac{\partial\vec{u}}{\partial t}+\vec{u}\cdot\nabla\vec{u}\right)=-\nabla p\,, \tag{1.60}$$

where $\vec{u}=\vec{u}(x,t)$. The remaining Eulerian equations of mass conservation and entropy advection follow from the constraints that s_0 and ϱ_0 are constant for fluid elements. Time differentiation and the transformations of (1.54) and (1.55) yield:

$$\frac{\partial\varrho}{\partial t}+\nabla\cdot(\varrho\vec{u})=0\,, \tag{1.61}$$

$$\frac{\partial s}{\partial t}+\vec{u}\cdot\nabla s=0\,. \tag{1.62}$$

Equations (1.60)–(1.62) together with a given function $e(\varrho,s)$ and the relation $p=\varrho^2\partial e/\partial\varrho=0$ constitute the Eulerian description.

(d) Ideal flow action principle in Eulerian coordinates

Early attempts to formulate an Eulerian version of Hamilton's principle (e.g. Clebsch [64] 1859, Bateman [23] 1929, Eckart [88] 1960) were ad hoc and only partly successful in that they yielded dynamical equations whose solutions are only a subset of the solutions to the ideal-fluid equations. The first general Eulerian versions of Hamilton's principle were those of Lin [162] (1963) and Seliger and Witham [236] (1968). For a thorough review of the early literature, see Finlayson [99] (1972). Lin's key contribution was the introduction of new constraints, which, as explained by Bretherton [49] (1970), force an equivalence between the Eulerian and particle-mechanics versions of Hamilton's principle. When the Eulerian version is derived from the particle-mechanics version, then Lin's constraints appear as automatic requirements, see Salmon [213].

Via (1.57) and the reasoning given above, the action functional must be equivalent to:

$$L=\int_{t_0}^{t_1}\int_\Omega\left[\frac{\vec{u}^2}{2}-e\left(\frac{\varrho_0(\vec{a})}{\det J(\vec{a},t)},s_0(\vec{a})\right)\right]\frac{\varrho_0(\vec{a})}{\det J(\vec{a},t)}\,\mathrm{d}^3x\,\mathrm{d}t\,, \tag{1.63}$$

for arbitrary variations $\delta\vec{a}$. To carry out this variation the entire integrand of (1.63) must be expressed in terms of $\vec{a}(\vec{x},t)$ and its derivatives. To express the velocity $\vec{u}$ as derivatives of $\vec{a}$, the identity:

$$\partial_t\vec{a}+(\vec{u}\cdot\nabla)\vec{a}=0\,, \tag{1.64}$$

is solved for $\vec{u}$ and the result substituted back into (1.63); equivalently, the vector equation (1.64) can be appended as constraints on (1.63) to allow for independent variations $\delta\vec{u}$ and $\delta\vec{a}$. If furthermore equation (1.61) is appended as a constraint on (1.63), then $\varrho_0/\det J$

can be replaced by ϱ and the following action functional:

$$L = \int_{t_0}^{t_1} \int_\Omega \left(\frac{\vec{u}^2}{2} - e(\varrho, s) - \zeta_i \mathrm{D}_t a_i \right) \varrho + \phi \left(\frac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho \vec{u}) \right) \mathrm{d}^3 x \, \mathrm{d}t \,, \tag{1.65}$$

is obtained for independent variations $\delta\varrho$, $\delta\vec{u}$, $\delta\vec{a}$, $\delta\vec{\zeta}$ and $\delta\phi$. Here $\vec{\zeta} = (\zeta_1, \zeta_2, \zeta_3)$ and ϕ are the Lagrange multipliers corresponding to (1.64) and (1.61). The velocity variation of (1.65) yields:

$$\delta\vec{u} : \ \vec{u} = \zeta_i \nabla a_i + \nabla\phi \,, \tag{1.66}$$

which can be used to eliminate $\vec{u}$ entirely. After integration by parts, (1.65) becomes:

$$L = \int_{t_0}^{t_1} \int_\Omega \varrho \left(\zeta_i \frac{\partial a_i}{\partial t} + \frac{\partial \phi}{\partial t} + \frac{\vec{u}^2}{2} + e(\varrho, s) \right) \mathrm{d}^3 x \, \mathrm{d}t \,, \tag{1.67}$$

for variations $\delta\vec{a}$, $\delta\vec{\zeta}$, $\delta\varrho$ and $\delta\phi$, in which $\vec{u}$ is simply an abbreviation for (1.66). The variational principle (1.67) is one of many obtained by Seliger and Witham [236] using a rather different approach; however, the above derivation emphasizes the close connection between (1.67) and the particle-mechanics form of Hamilton's principle, and it puts a clear physical interpretation on the "potentials" a_i as particle labels.

Variation of the Lagrangian (1.67) constitutes eight evolution equations for eight independent variables, while (1.63) suggests that the number of variables can be reduced with no loss in generality. Indeed, Boozer [41] gives a transparent proof that in the velocity representation (1.66) ϕ and ζ_3 can always be set to zero while for general non-homentropic flow, for instance, ζ_2 can be identified with the entropy s according to Salmon [213]. If the variable replacement $\vec{a} = (\beta, -\vartheta, \zeta)$, $\vec{\zeta} = (\alpha, s, 1)$ and $\phi = 0$ is employed, the more familiar version of the Lagrangian[9]:

$$L = \int_{t_0}^{t_1} \int_\Omega \ell(\zeta, \alpha, \beta, s, \vartheta, \varrho) \, \mathrm{d}^3 x \, \mathrm{d}t \,, \tag{1.68}$$

$$\ell = -\varrho \left[\frac{\partial \zeta}{\partial t} + \alpha \frac{\partial \beta}{\partial t} - s \frac{\partial \vartheta}{\partial t} + \frac{\vec{u}^2}{2} + e(\varrho, s) \right] , \tag{1.69}$$

is obtained for variations $\delta\zeta$, $\delta\alpha$, $\delta\beta$, δs, $\delta\vartheta$ and $\delta\varrho$, where:

$$\vec{u} = \nabla\zeta + \alpha\nabla\beta - s\nabla\vartheta \,. \tag{1.70}$$

[9] The variable replacement is employed to keep in line with [236]; correspondingly the minus sign in (1.69) is added without effect.

Variation of (1.68) yields the following Euler-Lagrange equations according to (1.47):

$$\delta\alpha:\ \mathrm{D}_t\beta = 0\,, \quad \delta\beta:\ \mathrm{D}_t\alpha = 0\,, \tag{1.71a}$$

$$\delta s:\ \mathrm{D}_t\vartheta = T\,, \quad \delta\vartheta:\ \mathrm{D}_t s = 0\,, \tag{1.71b}$$

$$\delta\zeta:\ \partial_t\varrho + \nabla\cdot(\varrho\vec{u}) = 0\,, \tag{1.71c}$$

$$\delta\varrho:\ \partial_t\zeta + \alpha\partial_t\beta - s\partial_t\vartheta + \vec{u}^2/2 + e + P/\varrho = 0\,. \tag{1.71d}$$

where temperature T is defined by (1.56). The Lagrangian (1.69) depends on the specific inner energy $e(\varrho, s)$, given in terms of ϱ, s, the three Clebsch potentials ζ, α, β and an additional potential field ϑ. The meaning of the latter becomes apparent by calculating the Euler Lagrange equation with respect to s, see (1.71b), giving the "potential representation":

$$\left\{\frac{\partial}{\partial t} + \vec{u}\cdot\nabla\right\}\vartheta = \mathrm{D}_t\vartheta = \frac{\partial e}{\partial s} = T\,, \tag{1.72}$$

for the temperature T, used three decades previously by Van Dantzig [262], who termed the field ϑ as the *thermasy*. Although still restricted to adiabatic and therefore reversible processes, the Lagrangian (1.69) represents a momentous step forward because of the rudimentary embedding of thermodynamics.

The statement (1.68), and slight modifications thereof, is the best-known Eulerian version of Hamilton's principle and it is straightforward to show that (1.70), (1.71) are equivalent to the ideal-fluid equations. However, (1.68), unlike (1.63), becomes generally invalid if the entropy is constant over even infinitesimally small volumes of the fluid. This failure occurs not for deep physical reasons, but because a locally constant entropy cannot serve as a particle label. The primary reason why (1.68) has achieved such popularity seems to be that it involves a minimal number of "non-physical" dependent variables $(\alpha, \beta, \vartheta)$. However, Salmon [213] argues that the "non-physical" variables are all either particle labels or closely related thereto and:

> *"[...] these labels would acquire an indisputable physical significance if only the internal energy were allowed to depend on solute concentrations that are conserved following the fluid particles. This is true whether or not the solute concentrations have topological properties that make them suitable themselves as particle-labelling variables."*

(e) Variational principles for viscous flow

As a result of the above considerations the interesting situation arises that variational principles exist when inertial terms are important and viscous terms are not and vice versa. Attempts to derive variational principles when both inertial and viscous terms

are included having so far failed. Rayleigh [200] noticed that the Helmholtz-Korteweg prinicple could be generalised to include the inertial term $\nabla(\vec{u}^2/2)$ in the equations, but not the full inertial term:

$$\vec{u} \cdot \nabla \vec{u} = \nabla(\vec{u}^2/2) - \vec{u} \times (\nabla \times \vec{u}) \,.$$

Millikan [178] gave the definitive treatment of the existence of a variational principle for the steady-state NS equations for an incompressible fluid by assuming a Lagrangian of the form:

$$\ell = \ell \left(\vec{u}, p, \partial \vec{u}/\partial t, \nabla \otimes \vec{u} \right) \,,$$

in terms of the velocity $\vec{u}$, the pressure p and their first order derivatives. By means of a detailed, lengthy argument he concluded that a variational principle could not be found unless $\vec{u} \cdot \nabla \vec{u} = 0$ or $\vec{u} \times (\nabla \times \vec{u}) = 0$. Later, Finlayson [98] derived the same result in a shorter and more concise form by the use of Fréchet differentials demonstrating that a variational principle for the NS equations cannot be obtained in the classical sense of Def. 1.1. In the same paper he also proposes an adjoint variational principle for the NS equations as an approach to the extended inverse problem according to Davis [80]. Other approaches in this direction following the developments described at the beginning of Sec. 1.2.2 are manifold [102, 146, 179, 264, 285] and not discussed further.

A general approach is clearly required based on the representation of the observable fields by potentials, i.e. by auxiliary fields representing the observables, similar to electrodynamics. A convincing explanation is given by Scholle [218] as to why in continuum theories the use of potentials is absolutely necessary for the construction of a Lagrangian: in order to fulfil the invariance with respect to the full Galilean group, at least one field must be non-measurable and therefore be a potential. In the same paper a general scheme for Lagrangians is constructed. Using Noether's theorem, canonical formulae give rise for identification of the relevant observable fields like mass density and flux density, momentum density, stress tensor, energy density and Poynting vector.

A combination of the variational principles of Seliger and Witham [236] and Helmholtz-Korteweg [124] is suggested by Scholle [216, 223] which involves the inertial terms and both shear and volume viscosity. The corresponding Lagrangian is shown to fulfil all the methodical requirements given in [218], however, by calculating the Euler-Lagrange equations a set of PDEs is induced which are, in the case of an incompressible flow, different from the original Navier-Stokes equations. The differences become manifest in a different form of the viscous terms, including their order (third order instead of second order terms), but also in an additional field, the thermasy, appearing explicitly. For compressible flow involving volume viscosity but with the shear viscosity neglected Zuckerwar and Ash [288]

obtained a similar potential-based Lagrangian. It is further noted that the mentioned approach is deeply related to a stochastic variational formulation by Arnaudon and Cruzeiro [13, 14], see also [148]. Furthermore, a comparison to Constantin et al. [71] (formula 2.12) – see also the work of Feireisl and Vasseur [94] related to a fluid mechanics model proposed by Brenner [47] – would be illuminating though not considered here.

1.3 Objectives and thesis structure

Approaches to the utilisation of potential fields in fluid mechanics are investigated linked to and building on the historical survey and literature review of the preceding section. An important feature being to demonstrate that the beneficial utilisation of auxiliary potential fields, which in the past has essentially been limited to ideal and Stokes-like flow cases, can be extended even to inertial *and* viscous flow, thus recovering parts of the advantages highlighted for classical electrodynamics in Sec. 1.1. In particular, the genesis of the two major strands followed is contained in Sec. 1.2.1 and 1.2.2: firstly the extension of the Eulerian flow action principles to the case of full NS flow, thus continuing the work of Clebsch [64], Helmholtz [124], Lin [162], Seliger and Witham [236], Finlayson [99] and others; secondly, generalisation of the idea to integrate the equations of motion along the lines of Clebsch [64] on the one hand, Sec. 1.2.1(a), and the lines of Legendre [159], Ranger [199], Coleman [68], Antanovskii [8] and more recently Scholle et al. [230] on the other hand, Sec. 1.2.1(b).

The thesis is comprised of two parts. Part I covers the underpinning theoretical foundations, at the end of which the options as to how to implement the ideas numerically are crystallised; their embodiment and application forming the foundation of Part II.

1.3.1 Part I

Part I builds on the historical theoretical underpinning described in Chapter 1 and encompasses the three different novel ideas reported in Chapters 2 - 3.

Chapter 2 is devoted to the problem of describing viscous flow in the framework of the Lagrange formalism. By comparing the potential representation (1.70) with the one proposed by Clebsch (1.23) for the isothermal case, it becomes apparent that any kind of extension of the system, by additional degrees of freedom as well as by additional physical effects, requires an adjustment of the potential representation, see e.g. Wagner [271]. Based on these preliminary findings, Scholle [222, 223] suggested a Lagrangian for viscous flow by suitably supplementing the Lagrangian (1.68) of [236] with additional terms from the Helmholtz principle (1.51), leading to partial success: the phenomenon “viscosity” occurs in a qualitatively correct manner but the equations of motion resulting from the

variation of Hamilton's principle differ from the NS equations significantly including their differential order and the occurrence of a further physical degree of freedom, the *thermasy*.

Following the ideas of [11, 288] it is proposed that the additional terms might be linked to effects *beyond thermodynamical equilibrium*, a hypothesis which is tested by means of several "benchmark problems"; these, however, reveal the occurrence of partly non-physical behaviour. As a remedy, a discontinuous Lagrangian is proposed which requires an abstract extension of the Lagrange formalism to discontinuities. Application of the latter leads to field equations which are discussed and interpreted in detail.

Closely related to the above new Lagrangian, which involves the Clebsch variables from the principle of ideal flow [236], is the question as to whether the Clebsch transformation itself can be extended beyond its classical inviscid flow context. The essential problem inhibiting the application of the Clebsch transformation to viscous flow is due to the friction force density $-\nu\Delta\vec{u}$ in the NS equations which does not seem to fit into the scheme (1.28). The problem of finding a decomposition of the form (1.28) with prescribed Clebsch variables α, β is handled for an *arbitrary* vector field $\vec{a}$ using a new approach. In consequence, a generalized Clebsch transformation is developed which also covers the case of incompressible viscous flow.

As discussed in Sec. 1.2.1 the integration approach to steady 2D-NS equations by Scholle et al. [230] is a promising continuation of the classical potential and Clebsch integration, and is also closely related to the complex variable method and Goursat approach originally known from linear elasticity. A major focus of **Chapter 3** is the extension of the above integration procedure to both the unsteady and 3D NS equations.

Initially, it is shown that the first integral formulation of steady 2D-NS flow can be extended to transient flow if the real-valued potential field Φ is replaced by an appropriately defined complex-valued one, i.e. an extension of the formulation necessarily goes together with the introduction of new auxiliary fields. Furthermore a similar first integral of the dynamic boundary condition, which is present in free surface applications, is established; the reformulated form, in the steady case, reduces to a pure Dirichlet-Neumann condition on the potential variable Φ constituting a key feature for the construction of efficient numerical solutions.

Based on the 2D tensor representation and analogy to Maxwell's equations the introduction of a tensor potential facilitates the NS equations being recast as the divergence of a tensor quantity set to zero. Integration leads to a tensor equation that splits conveniently into symmetric and skew-symmetric parts. Via the astute use of gauge freedoms, a decrease in the number of equations and unknowns is achieved as well as their transformation to a known, more tractable, equation set in which the differential order of the non-linear terms is reduced. Although consideration is focused on specific gauging of the

tensor potential, in order to ensure the equation set has a favourable structure, the theory itself is amenable to alternative development. Some of which offer the prospect of a promising continuation of the research field; for example, it is shown that the gauge freedoms can be utilised intelligently to establish a variational principle for steady viscous flow.

1.3.2 Part II

As an outcome of Chapter 3, amongst others, a major advantage of the first integral of the NS equations is the reduction of the associated non-linearity and the elegant treatment of the dynamic boundary condition when free surfaces are present, the latter applying essentially to steady flow. The construction of an efficient and competitive 2D-solver which in contrast to existing approaches [230] provides a general *generic* applicability and offers a convenient handling of free surfaces, reflects a first reasonable step to test the first integral approach and to identify numerical generalisations which are more suitable than the one utilised for demonstration purposes in Ch. 3. The development and application of such a solver forms the basis of Part II.

In **Chapter 4** the theoretical description of a new finite element method (FEM) based on the first integral of the NS equations is given, which is centred on a least-squares finite element (FE) framework requiring the set of equations to be of first order to allow for practical function spaces. Firstly, considerations are confined to the Stokes flow case. As a basis for a corresponding convergence analysis, a priori estimates for the analytic solution of the problem are necessary which, for a rather general class of elliptic boundary value problems, the present one included, can be obtained from the theory of Agmon, Douglis and Nirenberg (ADN) [2, 3]. The first integral system is classified as an elliptic system of Petrovskii type which proves to be particularly useful as it allows for an efficient least-squares minimisation with respect to the $L^2(\Omega)$-norm. The ellipticity of the corresponding bilinear form is demonstrated and the existence of a unique weak solution derived via the theorem of Lax-Milgram. The subsequent verification of optimal $H^1(\Omega)$-convergence mainly relies on the ADN-estimate and a standard interpolation argument for the associated FE spaces. Optimal $L^2(\Omega)$-convergence is technically more involved and utilises convergence and regularity results from the Galerkin discretisation of a second order auxiliary PDE system. Possible problems and an extension of the convergence analysis to non-smooth domains with corners are discussed and finally the fully non-linear problem is considered.

In **Chapter 5** implementation details for the fully non-linear problem involving an isoparametric concept of simplicial and quadrilateral elements of first and second order are provided. The method is thoroughly validated by checking the convergence rate of

different element and grid types for several test cases, and any potential problems with mass conservation and conditioning are addressed and resolved. The convergence of the non-linear iteration is compared to a standard formulation in primitive variables. One of the significant advantages of the new method is the beneficial structure of the resulting symmetric and positive definite systems thus avoiding the complication of having to solve saddle point problems. The solution of these systems is achieved in an efficient and scalable way using multigrid techniques: the adoption of a tailored algebraic multigrid (AMG) approach allows for fast and accurate solution of the linear systems resulting from large, sparse and *unstructured* grids.

Subsequent to deriving the above FEM including implementation details, the method is applied to several engineering problems of topical interest, from the fields of lubrication and film flow, in **Chapter 6**. These include the investigation and minimisation of friction between parallelly translating corrugated surfaces, material exchange between entrapped vortex regions and the bulk flow in a similar lubrication setting, and the case of gravity-driven film flow over corrugated surfaces. These applications not only demonstrate the flexibility and efficiency of the proposed method but also reveal new insights into these fields which are of interest in their own right.

Conclusions concerning the total body of work, together with recommendations representing potentially fruitful future avenues of research, are provided in **Chapter 7**.

Part I

New Potential Field Representations of the Navier-Stokes Equations

2 Approaches based on Clebsch-like representation

Since viscosity leads to dissipation and therefore to the irreversible transfer of mechanical energy to heat, a variational principle for viscous flow has to consider thermal degrees of freedom in order to remain consistent with Noether's theorem which implies conservation of energy for systems with time-translation invariance; otherwise the latter would have to be violated by an explicit time-dependence [216]. As recapitulated in Sec. 1.2.2(d), Seliger and Whitham [236] suggested how to embed thermal degrees of freedom in a variational formulation of fluid flow, via the Lagrangian (1.69) with the Clebsch-like velocity representation (1.70). By comparing the potential representation (1.70) with the one proposed by Clebsch (1.23) for the isothermal case, it is apparent that any kind of extension of the system, by additional degrees of freedom as well as by additional physical effects, requires an adjustment of the potential representation, see e.g. Wagner [271]. Scholle [218] provides a general explanation for the necessary use of different potential representations of the observables for different physical systems along the lines of a rigorous analysis of the fundamental symmetries the Lagrangian has to fulfil, with particular regard to Galilean invariance.

Based on these preliminary findings, Scholle [222, 223], along similar lines to Zuckerwar and Ash [288], suggested a Lagrangian for viscous flow by supplementing the Lagrangian given by (1.69) with additional terms leading to partial success: the phenomenon "viscosity" is captured in a qualitatively correct manner while the equations of motion resulting from the variation of Hamilton's principle differ from the NS equations noticeably; this is reviewed in detail below. Despite this progress, the need to improve the existing approach is obvious in order to obtain solutions from Hamilton's principle suitable for relevant flow problems. This is achieved based on an innovative idea by Anthony [11] involving a reformulation of the Lagrangian in terms of complex fields, which can also be understood as the inversion of Madelung's idea [168] of reformulating the complex Schrödinger's equation into a hydrodynamic form, and draws heavily on the work of Scholle [216].

2.1 Construction of the Lagrangian

First, as demonstrated in the prior work of [162, 222, 223, 271], Seliger and Whitham's Lagrangian (1.69) can be re-written alternatively as:

$$\ell = -\varrho \left[\mathrm{D}_t\zeta + \alpha \mathrm{D}_t\beta - s\mathrm{D}_t\vartheta - \frac{\vec{u}^{\,2}}{2} + e\left(\varrho, s\right) \right] , \tag{2.1}$$

in terms of the extended set of independent fields $\psi = (\vec{u}, \zeta, \alpha, \beta, \varrho, s, \vartheta)$ and their material time derivatives,

$$\mathrm{D}_t = \frac{\partial}{\partial t} + \vec{u} \cdot \nabla \, . \tag{2.2}$$

The above form of the Lagrangian yields two benefits: first, the potential representation (1.70) of the velocity field results from a variation with respect to $\vec{u}$ and hence does not need to be prescribed; second, by adding terms to the Lagrangian depending on first order derivatives of $\vec{u}$ in order to consider friction, the extended Lagrangian remains of first order. The latter is a useful feature because otherwise, i.e. in case of a Lagrangian containing second order derivatives of the fields, the computation of (i) the corresponding Euler-Lagrange equations and (ii) the canonical densities and flux densities resulting from Noether's theorem become more complicated. It is also very useful to avoid derivatives of order higher than one when applying Ritz's direct method to problems formulated in curvilinear coordinates.

2.1.1 Conventional approach and examples

Here the basic ideas and relevant findings of [216, 222, 223] are revisited in a concise form in order to capture the present state of the theory and identify the associated open problems.

In order to extend the Lagrangian (2.1) to incorporate viscosity it is reasonable to simply add further terms so as to modify the entropy balance:

$$-\frac{\partial}{\partial t}\left(\varrho s\right) - \nabla \cdot \left(\varrho s \vec{u}\right) = 0 \, , \tag{2.3}$$

which is obtained from (2.1) by variation with respect to the thermasy θ. The homogeneity of (2.3) indicates that only adiabatic processes are considered in (2.1). Note that the above balance alternatively results from Noether's theorem with respect to the transformation $\vartheta \to \vartheta + \varepsilon$ with $\varepsilon = \text{const}$, which is a symmetry transformation of the Lagrangian (2.1). In order to account for the production of entropy, this symmetry has to be broken which is easily achieved by adding a term linearly dependant on ϑ, i.e. $\vartheta\phi_d/T$, to the Lagrangian, where the dissipation heat, ϕ_d, should be positive and depend on the spatial derivatives of the velocity as the primary cause for the physical phenomenon "viscosity". Both are

satisfied by assuming ϕ_d has a quadratic dependence on $\partial_j u_i$, according to the classic literature on viscous flow [156, 241]. Finally, via the factor $1/T$ the character of the entropy as "weighted heat", according to $\delta Q = T\mathrm{d}S$, is accounted for.

The above considerations provided the motivation leading to the following extended Lagrangian [222, 223]:

$$\ell = -\varrho\left[\mathrm{D}_t\zeta + \alpha\mathrm{D}_t\beta - s\mathrm{D}_t\vartheta - \frac{\vec{u}^{\,2}}{2} + e\left(\varrho, s\right)\right] + \frac{\vartheta}{T}\left[\eta\mathrm{tr}\underline{D}^2 + \frac{\eta'}{2}\left(\nabla\cdot\vec{u}\right)^2\right], \tag{2.4}$$

in the absence of an external force and heat conduction; where η is the shear viscosity, η' the volume viscosity of the fluid and

$$\underline{D} := \frac{1}{2}\left[\nabla\otimes\vec{u} + \left(\nabla\otimes\vec{u}\right)^t\right], \tag{2.5}$$

is the shear rate tensor; tr denotes the trace of a tensor. The temperature T, according to classical thermodynamics, is given by (1.56). Now, by variation with respect to the thermasy ϑ, the following equation:

$$\frac{\partial}{\partial t}\left(\varrho s\right) + \nabla\cdot\left(\varrho s\vec{u}\right) = \frac{\eta}{T}\mathrm{tr}\underline{D}^2 + \frac{\eta'}{2T}\left(\nabla\cdot\vec{u}\right)^2, \tag{2.6}$$

results as an entropy balance with an entropy production rate on the right-hand side due to dissipation, as expected. Furthermore, the above Lagrangian fulfils the necessary symmetry requirements for Galilean invariance, as analysed in detail in Sec. A.1.1 of Appendix A. However, an unexpected feature arises: the momentum density $\vec{p}$, resulting as a canonical Noether observable, does not equal the mass flux density $\varrho\vec{u}$. The difference between both,

$$\vec{p}^{\,*} := \vec{p} - \varrho\vec{u} = -2\eta\nabla\cdot\left(\frac{\vartheta}{T}\underline{D}\right) - \eta'\nabla\left(\frac{\vartheta}{T}\nabla\cdot\vec{u}\right), \tag{2.7}$$

termed *quasi-momentum density*, needs to be explained physically. According to [218], $\vec{p}^{\,*}$ could be due to contributions to the system's momentum balance beyond the scope of the continuum hypothesis on a molecular scale, e.g. Brownian motion. This question is discussed in more detail in Sec. 2.4. Regardless, the dynamics induced by the Lagrangian (2.4) goes beyond the scope of classical theory: the resulting equations of motion differ significantly from the NS equations; in the case of incompressible flow and negligible buoyancy they read [216, 222, 223][1]:

$$\mathrm{D}_t\vec{u} = -\frac{\nabla p}{\varrho_0} + \nu\left\{\mathrm{D}_t + \nabla\otimes\vec{u}\right\}\left[2\underline{D}\nabla\left(\frac{\vartheta}{T}\right) + \frac{\vartheta}{T}\Delta\vec{u}\right] - \nu\mathrm{tr}\underline{D}^2\nabla\left(\frac{\vartheta}{T}\right), \tag{2.8}$$

[1] To allow a shorter notation in (2.8) and elsewhere, cf. (2.11) and (2.45), the material time derivative $\mathrm{D}_t[.]$ and the tensor $\nabla\otimes\vec{u}[.]$ are both interpreted as operators $\mathbb{R}^3\to\mathbb{R}^3$ and enclosed in curly parentheses; on the understanding that $\{\nabla\otimes\vec{u}\}\cdot[.]$ is interpreted as a matrix vector product.

$$\nabla \cdot \vec{u} = 0\,, \tag{2.9}$$

$$\mathrm{D}_t \left(\frac{\vartheta}{T} \right) = 1\,, \tag{2.10}$$

with constant mass density $\varrho = \varrho_0$ and kinematic viscosity $\nu := \eta/\varrho_0$. Looking at the set of PDEs (2.8)-(2.10), two striking features immediately become apparent, namely:

- the resulting field equations are third order PDEs, not second order ones like the NS equations;
- the thermasy ϑ appears explicitly in the field equations as an additional degree of freedom.

The above qualitative features have also been found in the case of compressible flow with pure volume viscosity by Zuckerwar and Ash [288], and the appearance of an additional physically relevant degree of freedom, in particular, appears also in the variational formulation of heat conduction proposed by Anthony [11], interpreted by him as a measure for the *deviation from the thermodynamical local equilibrium.* Similar assumptions are made in [288].

The temperature-weighted thermasy expression ϑ/T, which is a time quantity, together with its governing evolution equation (2.10) needs further elaboration. Firstly, note that a general solution of the linear equation (2.10) can be split into a particular and a homogeneous solution according to $\vartheta/T = t + \tau(\vec{x}, t)$ with $\mathrm{D}_t\tau = 0$ [216]; the particular solution indicating that in (2.8) may appear unusual terms which increase linearly with time. In order to guarantee a unique solution of the system (2.8)-(2.10) an initial condition for (2.10) is needed the prescription of which is not obvious at first sight and requires a physical interpretation.

However, since the thermasy is a potential (i.e. an auxiliary quantity giving the temperature as its material time derivative), a compelling physical meaning of the quantity itself as well as its evolution equation (2.10) is not obvious. Nevertheless, there are two speculative physical interpretations coming from van Dantzig [262] and Anthony [11]:

- In his classical paper, van Dantzig calls the quotient ϑ/T a "thermal time" indicating the time measured by a local observer moving in the fluid flow. Within this interpretation, the solution $\vartheta/T = t$, $\tau = 0$, is a distinguished candidate implying that all local "clocks" run synchronously, while other solutions with $\tau(\vec{x}, t) \neq 0$ imply asynchronous local clocks.
- Within the context of Anthony's approach $-\omega\vartheta/T$ is the phase of the complex field of thermal excitation. Therefore, its evolution equation describes the propagation of a kind of thermal wave. Also, according to Anthony, prescribing the initial condition

$\theta/T(\vec{x},0)=0$ would imply that the system is in thermodynamic equilibrium at $t=0$ which would not be the case for location-dependant initial conditions.

Although speculative, the above offer a plausible means of interpreting the additional terms and degrees of freedom in the evolution equations (2.6), (2.8)-(2.10) as an extension of the classical theory towards non-equilibrium thermodynamics. In order to test this hypothesis, the 'benchmark tests' (i)-(iii) are performed [222, 223], taking $\vartheta/T=t$ as a particular solution of the evolution equation (2.10), leading to the simplified form:

$$\mathrm{D}_t\vec{u}=-\frac{1}{\varrho_0}\nabla p+\nu\left\{\mathrm{D}_t+\nabla\otimes\vec{u}\right\}\left[t\Delta\vec{u}\right]\,, \tag{2.11}$$

of the equations (2.8). Thus only equations (2.9) and (2.11) remain to be considered:

(i) Plane Couette flow, i.e. shear flow between two parallel plates, the upper one of which is moving with a constant speed U, representing one of the most often encountered examples in fluid dynamics. Considering the unidirectional flow $\vec{u}=u(y)\vec{e}_x$ with boundary conditions $u(0)=0$ and $u(h)=U$, the solution of (2.9), (2.11) is given by the linear velocity profile $u(y)=Uy/h$ and a constant pressure p in full accordance with the NS equations.

(ii) As an example of a transient flow, the sudden movement of a flat plate is considered: a horizontal plate of infinite extent is covered by a fluid at rest at $t=0$, see Fig. 2.1a. At $t=0$ the plate suddenly starts moving from rest ($u(0)=0$) with constant velocity U in the horizontal direction, invoking a flow within the fluid. Since no characteristic length is contained in this problem, a representation of the velocity profile in terms of a similarity variable, $\vec{u}=Uf(\xi)$ with $\xi=y/\sqrt{\nu t}$ (boundary conditions $f(0)=1$, $f(\infty)=0$) [241], is used leading to the solution $f(\xi)=\exp(-\xi)$ of the field equations (2.9), (2.11), whereas the classical solution of the NS equations reads $f(\xi)=1-\mathrm{erf}\,(\xi/2)$ [241] involving the error function erf. In Fig. 2.1a both resulting profiles are compared to each other, revealing qualitatively concordant profiles with quantitative differences.

(iii) The Lamb-Oseen vortex [154], the flow geometry of which is shown in Fig. 2.1b, is another example of a transient flow: using cylindrical coordinates r,φ,z, a similarity variable $\xi:=r/\sqrt{\nu t}$ is available allowing for a representation of the velocity as $\vec{u}=u(r,t)\vec{e}_\varphi$ with the time-dependant profile $u(r,t)=\Gamma f\left(\xi\right)/(2\pi r)$ for a given circulation Γ. Using boundary conditions $f(\infty)=1$, $f(0)=0$, from (2.9), (2.11) the analytical form $f(\xi)=1-\xi K_1(\xi)$ is obtained with the modified Bessel function of first order, K_1 [216, 222, 223]. In contrast, $f(\xi)=1-\exp(-\xi^2/4)$ is obtained in the classical solution [154]. Both solutions are compared in Fig. 2.1b, revealing again qualitative accordance with quantitative differences.

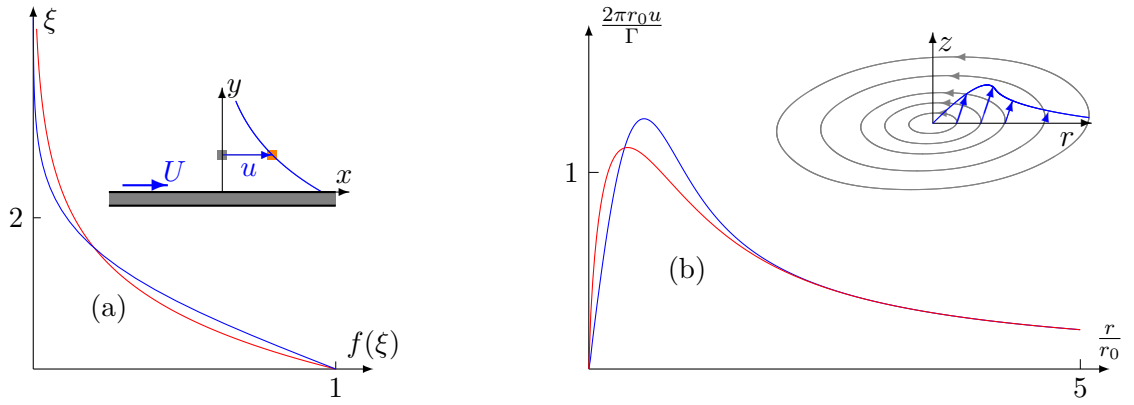


Figure 2.1: Resulting velocity profiles for (a) the flow generated by a plate that is suddenly moved and (b) the Lamb-Oseen vortex (at time $t = \pi\varrho r_0^2/25\eta$). For both examples the solution resulting from the field equations (2.9), (2.11) (red) are compared to the respective profile resulting from the original NS equations (blue).

In summary, it can be ascertained from the above examples that the phenomenon of viscosity is at least captured qualitatively by Hamilton's principle. Hence, the Lagrangian (2.4) reflects a relevant step forward to a satisfying description of viscous flow within the framework of Lagrange formalism. On the one hand, the differences between the classical theory and the results obtained by variation are quite pronounced for the transient flow examples (ii) and (iii); it is open to dispute if they can be explained as non-equilibrium effects. On the other hand, alternative equations of motion containing additional terms related to thermal relaxation are discussed in the literature; for instance in the work of Ash and Zardadkhan [16, 17] who make use of the NS equations supplemented with a 'pressure relaxation term', implying a vortex solution considerably different from the classical Lamb-Oseen vortex, in order to resolve discrepancies between existing models and observations for dust devils and tornados.

Three new additional examples are explored in order to further address the question of whether the non-classical form of the field equations (2.8)-(2.10) can consistently be explained due to non-equilibrium thermodynamics. Note that in these examples additional solutions to that of (2.10) (namely, $\vartheta/T \neq t$) are considered:

(iv) Taylor-Couette flow between two cylinders of radius r_i and r_a, see Fig. 2.2a, invoked by rotation of the inner cylinder with angular velocity ω_0, is characterised by a velocity profile $\vec{u} = u(r)\vec{e}_\varphi$. By applying this to equations (2.9), (2.11) with boundary conditions $u(r_i) = \omega_0 r_i$ and $u(r_a) = 0$, the following solution is obtained:

$$u(r) = \frac{\omega_0 r_i^2 r_a}{r_a^2 - r_i^2}\left[\frac{r_a}{r} - \frac{r}{r_a}\right],$$

$$p(r) = \varrho \int \frac{u(r)^2}{r} \mathrm{d}r \,,$$

in perfect agreement with classical theory.

(v) Plane Poiseuille flow between two parallel plates a distance h apart, driven according to Fig. 2.2b by a pressure gradient $(p_1 - p_2)/l$, for which a unidirectional velocity field $\vec{u} = u(y)\vec{e}_x$ is assumed. Considering no-slip conditions $u(0) = u(h) = 0$ at the lower and upper plate, the solution of (2.9), (2.11) is given by:

$$u(y) = Ky(h - y)/2 \,, \tag{2.12}$$

$$p(x, y, t) = p_1 + \eta K \left[x + u(y)t\right] \,. \tag{2.13}$$

By identifying K to be $(p_1 - p_2)/(\eta l)$, the above velocity profile, (2.12), is in perfect agreement with the classical solution [241]. However, the associated pressure, (2.13), contains an additional term $\eta K u(y)t$, by which the adherence of the boundary conditions for the pressure p at the inflow and the outflow is inhibited. Moreover, the pressure is unsteady and, as a non-physical feature, it tends to infinity with increasing time. The reason for the latter problem stems from the choice of the particular thermasy solution $\vartheta/T = t$ of the evolution equation (2.10) which increases with time; a fact that seems to be problematic not only for this specific flow problem but for problems in fluid mechanics in general, as stated in [220]. In their response [287] to the comment [220], Zuckerwar and Ash suggested constructing a time-independent solution of the evolution equation (2.10), fulfilling $\vec{u} \cdot \nabla(\vartheta/T) = 1$. Following their suggestion, as a steady solution for the weighted thermasy the expression:

$$\frac{\vartheta}{T} = \frac{x}{u(y)} + f_1(y) \,, \tag{2.14}$$

is obtained with arbitrary integration function $f_1(y)$, see Appendix A.1.2. The associated solution of (2.8), (2.9) reads, accordingly, in implicit form as:

$$y = \sqrt{\frac{2u}{K}}\, {}_2F_1\left(\frac{1}{6}, \frac{1}{2}; \frac{7}{6}; u^3/u_{\max}^3\right) \,, \tag{2.15}$$

$$u_{\max} = \frac{9\Gamma(\frac{2}{3})^2\Gamma(\frac{5}{6})^2}{8\pi^3} K h^2 \,, \tag{2.16}$$

$$p = p_0 - \eta K x + \eta \frac{\vartheta}{T} \frac{u'(y)^2}{2} \,, \tag{2.17}$$

with the Gaussian hypergeometric function ${}_2F_1$ and a constant of integration p_0. The velocity profile given by (2.15), (2.16) is visualised in Fig. 2.2b (right) and differs markedly from the classical parabolic profile (2.12) (left), due to the fact that its first order derivative vanishes at the walls, $u'(0) = u'(h) = 0$, thus indicating

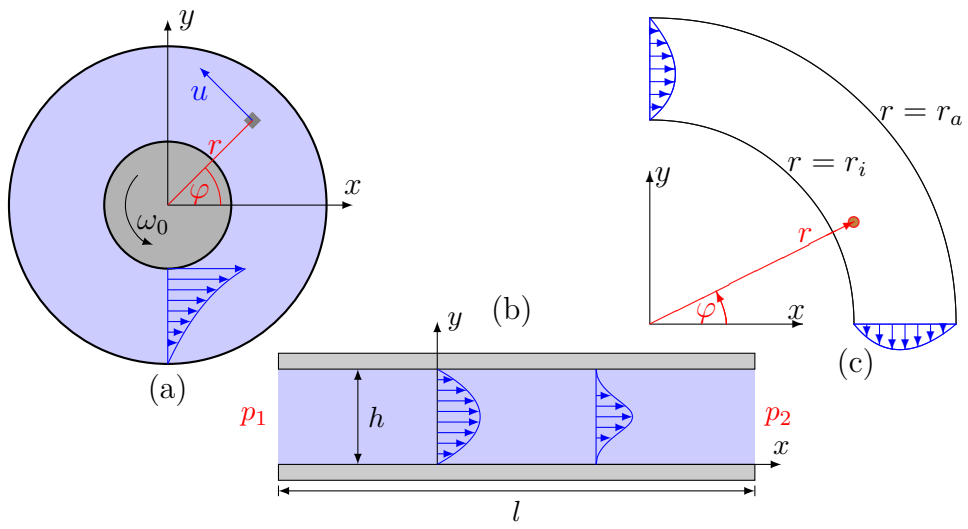


Figure 2.2: Flow geometry and velocity profiles of problems (iv)-(vi): (a) Taylor-Couette flow, (b) plane Poiseuille flow and (c) Poiseuille flow in a curved channel.

a zero wall shear stress. As another conspicuous feature, the y-dependence of the pressure inhibits again fulfilment of the boundary conditions for the pressure p at the inflow and the outflow.

The above shows, it is not possible to recover the classical steady solution for plane Poisseuille flow, based on the system (2.8)-(2.10). By setting $\vartheta/T = t$ the correct velocity profile is recovered but a non-physical, time-growing pressure is obtained; by writing $\vec{u} \cdot \nabla(\vartheta/T) = 1$ as suggested by [287], the problem of the explicit time-dependency is removed but the velocity profile is non-physical. In both cases a y-dependant pressure inhibits the fulfilment of the pressure boundary conditions. Although not proven, it is unlikely that *any* solution of (2.10) can be expected which overcomes both the explicit time and spatial dependency.

(vi) Poiseuille flow in a curved channel, see Fig. 2.2c, was investigated by Richter [203] who discovered qualitatively the same problems associated with plane Poiseuille flow. In particular, the resulting pressure solution,

$$p = 2\frac{p_2 - p_1}{\pi}\left[\varphi + \frac{u(r)t}{r}\right] + \varrho \int \frac{u(r)^2}{r} \mathrm{d}r\,, \tag{2.18}$$

again contains a non-physical term increasing with time. Other attempts to find a time-independent solution matching the boundary conditions have remained unsuccessful as in the case of the previous example.

Summarising the above benchmark tests, for only two of the six examples, namely (i) and (iv), is the classical solution recovered. For the two transient flows, (ii) and (iii), "reasonable" solutions are obtained, exhibiting quantitatively different velocity profiles compared to the classical ones. For the two pressure-driven flows (v) and (vi) no adequate solutions of the field equations can be constructed which simultaneously fulfil the pressure boundary conditions. Hence, the variational principle based on the Lagrangian (2.4) does not recover the dynamics of viscous flow in a proper way, since its applicability seems restricted to special flow problems only.

Moreover, four of the six benchmark solutions contradict the hypothesis that the differences compared to classical theory can be explained by effects beyond the scope of thermodynamic equilibrium: apart from the non-physical features discovered above, one would expect that the non-equilibrium solution tends towards the classical equilibrium solution if a special relaxation parameter in the problem, physically related to the deviation from equilibrium, tends to infinity. This is not the case here since no additional parameters exist but mass density, viscosity and specific heat.

Analysing the above examples in more detail, the explicit appearance of the weighted thermasy ϑ/T in the equations of motion (2.8) seems to be the crucial issue leading to non-physical solutions, since ϑ/T turns out to have unlimited growth, either spatially or temporally, which also prohibits its interpretation in connection with non-equilibrium thermodynamics. Moreover, the anomalous relation (2.7) between mass flux density and momentum density implies that the discrepancy between mass flux and momentum also tends to increase spatially or temporally due to the ϑ/T-dependence.

In line with the aims of this Chapter, a modified form of the Lagrangian (2.4) is developed below, as published in Scholle and Marner [226], which overcomes these anomalies.

2.1.2 Constructing a non-conventional Lagrangian based on complex fields

In 1927 Madelung [168] discovered a remarkable analogy between quantum mechanics and fluid mechanics by reformulating the complex Schrödinger's equation into a hydrodynamic form: by decomposing the quantum mechanical state function ψ into modulus and phase according to:

$$\psi = \sqrt{\frac{\varrho}{m}} \exp\left(\mathrm{i}\frac{m}{\hbar}\phi\right) , \tag{2.19}$$

Schrödinger's equation is transformed to a set of PDEs[2]:

$$\frac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho \nabla \phi) = 0 , \tag{2.20}$$

[2]Cf. the considerations in the introduction, Sec. 1.1: equations (2.20), (2.21) are special cases of the more general equations (1.15), (1.17) which additionally include the vector potential and thus imply a velocity field (1.16) which admits non-zero vorticity.

$$\frac{\partial\phi}{\partial t} + \frac{1}{2}\left(\nabla\phi\right)^2 - \frac{\hbar^2\Delta\sqrt{\varrho}}{2m^2\sqrt{\varrho}} + \frac{U}{m} = 0\,. \tag{2.21}$$

These are obviously the equations of motion for a kind of fluid, the so-called *Madelung fluid*: equation (2.20) is the continuity equation and (2.21) is Bernoulli's equation for a fluid with the "unusual" pressure function $P = -\hbar^2\Delta\sqrt{\varrho}/(2m^2\sqrt{\varrho})$ and with vorticity-free velocity field $\vec{u} = \nabla\phi$. Based on these substitutions, Madelung established a fluid mechanics picture of Schrödinger's theory.

Many years later Anthony [11] suggested the inversion of this idea, i.e. a "Schrödinger-picture" of fluid mechanics and thermodynamics by combining the density ϱ and the Clebsch variable ζ in (1.69) to form a complex *matter field* ψ according to (2.19). Moreover, he introduced two more complex fields, namely a complex vortex potential Ω by combining the two remaining Clebsch variables α, β and the complex field of thermal excitation χ, giving the temperature as its absolute square: $T = \bar{\chi}\chi$. The motivation for this transformation is originally given by Anthony's entropy concept: the entropy balance results from a canonical procedure related to the phase translation invariance of the complex fields as a balance of second kind within the framework of second variation and related stability criteria, see [216] for further details. Additionally, Anthony states that by the complex representation a basic concept is given for an accurate formulation of thermodynamics of irreversible processes within the framework of Lagrange formalism.

For convenience only a partial transformation to complex fields is applied in a slightly modified form to the Lagrangian (2.4): by introducing T_0 as a constant reference temperature and c_0 as a reference constant for the specific heat and considering the identity:

$$-s\mathrm{D}_t\vartheta = \mathrm{D}_t\left[(c_0 - s)\vartheta\right] - c_0T_0\exp\left(\frac{s}{c_0}\right)\mathrm{D}_t\left[\exp\left(-\frac{s}{c_0}\right)\frac{\vartheta}{T_0}\right]\,,$$

motivation is provided for a generalised definition of the field of thermal excitation as:

$$\chi := \sqrt{c_0T_0}\exp\left(\frac{s}{2c_0} - \mathrm{i}\omega_0\exp\left(-\frac{s}{c_0}\right)\frac{\vartheta}{T_0}\right)\,, \tag{2.22}$$

supplemented by the substitution:

$$\phi := \zeta + (c_0 - s)\vartheta\,, \tag{2.23}$$

for the Clebsch variable ζ. Note that in (2.22) another constant, ω_0, is introduced due to dimensional reasons, like T_0 and c_0 before. Although there is no general rule as to how the three constants T_0, c_0 and ω_0 have to be chosen, it is reasonable to choose the reference temperature T_0 as a "typical" temperature and c_0 as a "typical" specific heat. In the case of an incompressible flow with constant specific heat discussed subsequently, Subsec. 2.1.3, the choice of c_0 is obvious and the choice of T_0 arbitrary since the resulting

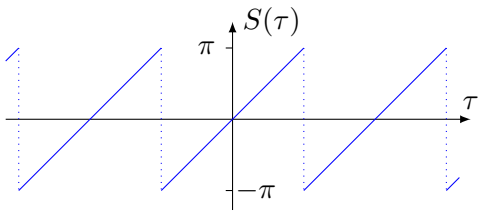


Figure 2.3: The sawtooth function.

Lagrangian (2.32) does not depend on it any more. In contrast, the choice of ω_0 is not obvious or how the physics is affected by it. This is analysed and discussed carefully in the following.

From the above substitutions (2.22), (2.23), it follows that:

$$\mathrm{D}_t\zeta - s\mathrm{D}_t\vartheta = \mathrm{D}_t\phi + \frac{1}{\omega_0}\mathrm{Im}\left(\bar{\chi}\mathrm{D}_t\chi\right) , \tag{2.24}$$

$$S\left(\omega_0 \exp\left(-\frac{s}{c_0}\right)\frac{\vartheta}{T_0}\right) = -\arg\chi = -\mathrm{i}\ln\sqrt{\frac{\bar{\chi}}{\chi}} , \tag{2.25}$$

with the sawtooth function (see also Fig. 2.3):

$$S(x) := x - 2\pi\left\lfloor\frac{x+\pi}{2\pi}\right\rfloor . \tag{2.26}$$

While the first equation (2.24) allows for a unique transformation of the respective real-valued terms in the Lagrangian (2.4) in terms of the complex thermal excitation field, the second equation (2.25) reveals that no equivalent for the thermasy ϑ explicitly appearing in the friction term of (2.4) can be constructed in terms of the complex field χ. The reason for this is the non-uniqueness of the argument of a complex number. The obvious and most feasible way to resolve this issue is the use of:

$$\frac{T_0}{\mathrm{i}\omega_0}\exp\left(\frac{s}{c_0}\right)\ln\sqrt{\frac{\bar{\chi}}{\chi}} = \frac{\bar{\chi}\chi}{\mathrm{i}\omega_0 c_0}\ln\sqrt{\frac{\bar{\chi}}{\chi}} , \tag{2.27}$$

as a substitute for ϑ, leading to the modified Lagrangian:

$$\ell = -\varrho\left[\mathrm{D}_t\phi + \alpha\mathrm{D}_t\beta + \frac{1}{\omega_0}\mathrm{Im}\left(\bar{\chi}\mathrm{D}_t\chi\right) - \frac{\vec{u}^{\,2}}{2} + e\right] + \frac{\bar{\chi}\chi\ln\sqrt{\frac{\bar{\chi}}{\chi}}}{\mathrm{i}\omega_0 c_0 T}\left[\eta\mathrm{tr}\underline{D}^2 + \frac{\eta'}{2}\left(\nabla\cdot\vec{u}\right)^2\right] . \tag{2.28}$$

Comparing this Lagrangian with (2.4), two obvious differences are discernible: first, (2.28) is *discontinuous* due to the logarithmic term; second, the angular frequency ω_0, which has primarily been introduced for dimensional reasons, becomes a relevant parameter,

the physical meaning of which will be clarified subsequently. The most striking feature, however, is that the unlimited weighted thermasy ϑ/T appearing in (2.4) has been replaced by an expression with limited values between $-\pi$ and π.

2.1.3 Incompressible flow with constant specific heat and external force

Consider a fluid flow with constant mass density, constant specific heat and without thermal expansion:

$$\varrho = \varrho_0 \,, \tag{2.29}$$

$$s = c_0 \ln\left(\frac{T}{T_0}\right) \,, \tag{2.30}$$

$$e = c_0 T \,. \tag{2.31}$$

Note that (2.30) implies $\bar{\chi}\chi = c_0 T$ for the field of thermal excitation (2.22) in accordance with Anthony [11]. Furthermore, since for the incompressible case volume viscosity is excluded from the very beginning, the term involving η' in (2.28) is absent. Finally, adding the specific potential energy $V = V(\vec{x}, t)$ of an external force, leads to the simplified Lagrangian:

$$\ell = -\varrho_0 \left[\mathrm{D}_t\phi + \alpha \mathrm{D}_t\beta + \frac{1}{\omega_0}\mathrm{Im}\left(\bar{\chi}\mathrm{D}_t\chi\right) - \frac{\vec{u}^{\,2}}{2} + \bar{\chi}\chi + V - \frac{\nu}{\mathrm{i}\omega_0}\ln\sqrt{\frac{\bar{\chi}}{\chi}}\mathrm{tr}\underline{D}^2 \right] . \tag{2.32}$$

In conventional variational calculus, Euler-Lagrange equations can be computed if the Lagrangian is two times continuously differentiable [108]. If this basic requirement is not fulfilled, as in (2.32), a non-standard approach is required for variation, which is developed in the following section.

2.2 Variation of the discontinuous Lagrangian: general formalism

Consider a variational principle $\delta I = 0$ where I is given by:

$$I = \int_{t_1}^{t_2} \iiint_V \ell\left(\psi_i, \dot{\psi}_i, \nabla\psi_i\right) \mathrm{d}V\mathrm{d}t \,, \tag{2.33}$$

depending on independent fields ψ_i with $i = 1, \cdots, N$ and $\psi_N = \varphi$. The Lagrangian ℓ is assumed to be discontinuous with respect to φ at fixed values φ_n, $n = 1, \cdots, N_S$, but continuously differentiable with respect to all other fields, ψ_i with $i < N$, and also continuously differentiable with respect to all derivatives of any field, including $\dot{\varphi}$ and $\nabla\varphi$.

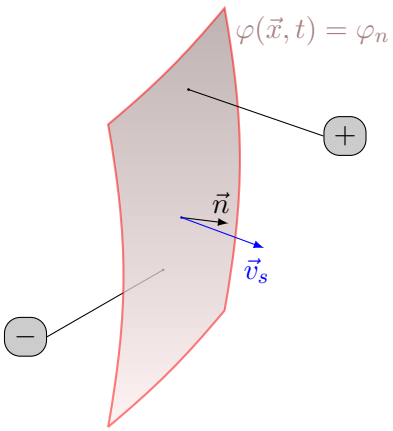


Figure 2.4: Surface S_n, along which a discontinuity becomes manifest.

In 3D space, the discontinuities with respect to φ become manifest along surfaces $S_n(t)$ defined by:

$$S_n := \{\vec{x}\,|\,\varphi(\vec{x},t)=\varphi_n\}\;, \quad n=1,\cdots,N_S\,, \tag{2.34}$$

intersecting the system's volume V into a finite number, N_S+1, of sub-volumes according to:

$$V=\sum_{n=0}^{N_S} V_n\,. \tag{2.35}$$

The sub-volume V_n denotes the region between S_n and S_{n+1} other than V_0 and V_{N_S} which denote the region between the system's boundary ∂V and S_1 or S_{N_S}, respectively.

From a physical viewpoint, these time-dependent interfaces, S_n, can be related to any kind of discontinuous phenomena such as phase boundaries between non-mixable fluids, propagating shock fronts in gaseous media or flame fronts. Their local propagation velocity is denoted by $\vec{v}_s$ and $\vec{n}$ is the normal vector of the interface, see Fig. 2.4. The orientation of $\vec{n}$ is defined by the convention $\vec{n}\cdot\vec{v}_s>0$.

2.2.1 Euler-Lagrange equations

First, only the subset of variations $\delta\psi_i=0$ is considered with $\delta\psi_i=0$ at the interfaces S_n and at the system's boundary ∂V. Free variation is assumed inside the sub-volumes V_n. Note, that this kind of variation does not cause any shift of the interfaces S_n. Under these assumptions, the usual derivation procedure leading to the Euler-Lagrange equations can be performed separately inside each sub-volume and hence, the wellknown Euler-Lagrange

equations [108]:

$$\mathrm{EL}_i := \frac{\partial \ell}{\partial \psi_i} - \frac{\partial}{\partial t}\left(\frac{\partial \ell}{\partial \dot{\psi}_i}\right) - \nabla \cdot \left(\frac{\partial \ell}{\partial \nabla \psi_i}\right) = 0\,, \tag{2.36}$$

remain valid piecewise at each sub-volume V_n.

2.2.2 Matching conditions

Next, a larger set of variations with $\delta\psi_i \neq 0$ at the interfaces S_n is considered for $i = 1, \cdots, N-1$, but the variation of $\psi_N = \varphi$ is again restricted to $\delta\varphi = 0$ in order to exclude any shift of the interfaces themselves. As before, $\delta\psi_i = 0$ is again prescribed at the system's boundary. Now, variation with respect to $\psi_1, \cdots, \psi_{N-1}$ leads to:

$$\begin{aligned}
\delta I &= \int_{t_1}^{t_2} \sum_{n=0}^{N_S} \iiint_{V_n} \left[\frac{\partial \ell}{\partial \psi_i}\delta\psi_i + \frac{\partial \ell}{\partial \dot{\psi}_i}\delta\dot{\psi}_i + \frac{\partial \ell}{\partial \nabla \psi_i}\nabla\delta\psi_i\right] \mathrm{d}V\,\mathrm{d}t\,, \\
&= \int_{t_1}^{t_2} \sum_{n=0}^{N_S} \iiint_{V_n} \left[\mathrm{EL}_i\,\delta\psi_i + \frac{\partial}{\partial t}\left(\frac{\partial \ell}{\partial \dot{\psi}_i}\delta\psi_i\right) + \nabla \cdot \left(\frac{\partial \ell}{\partial \nabla \psi_i}\delta\psi_i\right)\right] \mathrm{d}V\,\mathrm{d}t\,,
\end{aligned} \tag{2.37}$$

using the abbreviation (2.36) for the Euler-Lagrange expressions. By means of Gauss's theorem:

$$\iiint_{V_n} \nabla \cdot \left(\frac{\partial \ell}{\partial \nabla \psi_i}\delta\psi_i\right) \mathrm{d}V = \oiint_{\partial V_n} \vec{n} \cdot \frac{\partial \ell}{\partial \nabla \psi_i}\delta\psi_i\,\mathrm{d}S\,,$$

and Reynolds' transport theorem well-known from fluid dynamics [241]:

$$\iiint_{V_n} \frac{\partial}{\partial t}\left(\frac{\partial \ell}{\partial \dot{\psi}_i}\delta\psi_i\right) \mathrm{d}V = \frac{\mathrm{d}}{\mathrm{d}t} \iiint_{V_n} \frac{\partial \ell}{\partial \dot{\psi}_i}\delta\psi_i\,\mathrm{d}V - \oiint_{\partial V_n} \vec{n} \cdot \vec{v}_s \frac{\partial \ell}{\partial \dot{\psi}_i}\delta\psi_i\,\mathrm{d}S\,,$$

with $\vec{v}_s$ being the velocity of the propagating interface S_n, see Fig. 2.4, the variation takes the form:

$$\begin{aligned}
\delta I = &\int_{t_1}^{t_2} \sum_{n=0}^{N_S} \iiint_{V_n} \mathrm{EL}_i\,\delta\psi_i\,\mathrm{d}V\mathrm{d}t + \sum_{n=0}^{N_S} \iiint_{V_n} \frac{\partial \ell}{\partial \dot{\psi}_i}\delta\psi_i\,\mathrm{d}V\Bigg|_{t_1}^{t_2} \\
&+ \int_{t_1}^{t_2} \sum_{n=0}^{N_S} \oiint_{\partial V_n} \vec{n} \cdot \left[\frac{\partial \ell}{\partial \nabla \psi_i} - \vec{v}_s \frac{\partial \ell}{\partial \dot{\psi}_i}\right] \delta\psi_i\,\mathrm{d}S\mathrm{d}t\,.
\end{aligned} \tag{2.38}$$

Since $\delta\psi_i = 0$ at the initial and final time $t_{1,2}$, the above identity simplifies, in line with the Euler-Lagrange equations (2.36), to:

$$\delta I = \int_{t_1}^{t_2} \sum_{n=0}^{N_S} \oiint_{\partial V_n} \vec{n} \cdot \left[\frac{\partial \ell}{\partial \nabla \psi_i} - \vec{v}_s \frac{\partial \ell}{\partial \dot{\psi}_i} \right] \delta\psi_i \, \mathrm{d}S \mathrm{d}t \,.$$

In the following the limit of the respective discontinuous expression by approaching it from the front side (subscript $+$) or the rear side (subscript $-$) of the interface S_n is indicated by $[\cdots]_\pm$. The front and rear side are defined by the orientation of the normal vector $\vec{n}$ as in Fig. 2.4. Then, by decomposing each surface integral over ∂V_n in one integral along the front side of S_n and another one along the rear side of S_{n+1}, in general:

$$\begin{aligned}
\sum_{n=0}^{N_S} \oiint_{\partial V_n} \vec{n} \cdot [\cdots] \, \delta\psi_i \, \mathrm{d}S = \sum_{n=1}^{N_S-1} \left[\iint_{S_{n+1}} \vec{n} \cdot [\cdots]_- \, \delta\psi_i \, \mathrm{d}S - \iint_{S_n} \vec{n} \cdot [\cdots]_+ \, \delta\psi_i \, \mathrm{d}S \right] \\
+ \iint_{S_1} \vec{n} \cdot [\cdots]_- \, \delta\psi_i \, \mathrm{d}S - \iint_{S_{N_S}} \vec{n} \cdot [\cdots]_+ \, \delta\psi_i \, \mathrm{d}S \\
= \sum_{n=1}^{N_S} \iint_{S_n} \vec{n} \cdot \Big([\cdots]_- - [\cdots]_+ \Big) \, \delta\psi_i \, \mathrm{d}S \,, \qquad (2.39)
\end{aligned}$$

and, in particular:

$$\delta I = \int_{t_1}^{t_2} \sum_{n=1}^{N_S} \iint_{S_n} \vec{n} \cdot \left[\left[\frac{\partial \ell}{\partial \nabla \psi_i} - \vec{v}_s \frac{\partial \ell}{\partial \dot{\psi}_i} \right] \right] \delta\psi_i \, \mathrm{d}S \mathrm{d}t \,,$$

where the double square bracket indicates the jump at the interface: $[[\ldots]] := [\cdots]_- - [\cdots]_+$. Thus, variation $\delta I = 0$ leads to:

$$\vec{n} \cdot \left[\left[\frac{\partial \ell}{\partial \nabla \psi_i} - \vec{v}_s \frac{\partial \ell}{\partial \dot{\psi}_i} \right] \right] = 0 \,, \qquad (2.40)$$

as *matching conditions* for the generalized fluxes, $i = 1, \cdots, N-1$ at each interface.

Independent of the formal proof given above the matching conditions can also be understood as *natural boundary conditions* at the phase boundaries in a multiphase flow when assuming that all phases of the flow consist of the same liquid, leading to the same equation (2.40).

2.2.3 Production condition

At the system's boundary $\delta\psi_i = 0$ is prescribed; apart from this, free variation of all fields is allowed everywhere inside V, including free variation of $\psi_N = \varphi$. As a consequence

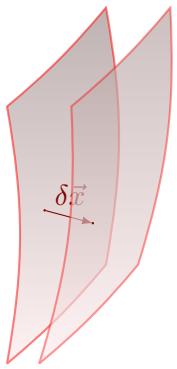


Figure 2.5: Variation of an interface caused by variation of φ.

of the latter, the position of the interfaces S_n is varied, too: an arbitrary point $\vec{x}$ of S_n defined by (2.34) is shifted to a different position $\vec{x}+\delta\vec{x}$ according to:

$$\begin{aligned}\varphi_n &= \varphi\left(\vec{x}+\delta\vec{x},t\right)+\delta\varphi\left(\vec{x}+\delta\vec{x},t\right)=\varphi\left(\vec{x},t\right)+\delta\vec{x}\cdot\nabla\varphi\left(\vec{x},t\right)+\delta\varphi\left(\vec{x},t\right)+\mathcal{O}\left(\delta\vec{x}^2\right)\\ &= \varphi_n+\delta\vec{x}\cdot\nabla\varphi\left(\vec{x},t\right)+\delta\varphi(x,t)+\mathcal{O}\left(\delta\vec{x}^2\right)\,,\end{aligned}$$

leading to the identity:

$$\delta\vec{x}\cdot\nabla\varphi=-\delta\varphi\,, \tag{2.41}$$

for the local shift $\delta\vec{x}$ of the discontinuity interface S_n. In the case that S_n is shifted in the forward direction, $\vec{n}\cdot\delta\vec{x}>0$, in a thin layer of thickness $\vec{n}\cdot\delta\vec{x}$ the value of ℓ is changed from $[\ell]_+$ to $[\ell]_-$; hence the integral (2.37) has to be supplemented by the following correction term:

$$\int_{t_1}^{t_2}\sum_{n=1}^{N_S}\iint_{S_n}[[\ell]]\,\vec{n}\cdot\delta\vec{x}\,\mathrm{d}S\mathrm{d}t\,.$$

Since $\nabla\varphi\parallel\vec{n}$, the identity $\vec{n}\delta\varphi=-(\delta\vec{x}\cdot\nabla\varphi)\vec{n}=-(\delta\vec{x}\cdot\vec{n})\nabla\varphi$ results, leading to:

$$\delta I=\int_{t_1}^{t_2}\sum_{n=1}^{N_S}\iint_{S_n}\left(-\nabla\varphi\cdot\left[\left[\frac{\partial\ell}{\partial\nabla\varphi}-\vec{v}_s\frac{\partial\ell}{\partial\dot{\varphi}}\right]\right]+[[\ell]]\right)\vec{n}\cdot\delta\vec{x}\,\mathrm{d}S\mathrm{d}t\,.$$

Variation with respect to φ leads to the jump condition:

$$\nabla\varphi\cdot\left[\left[\frac{\partial\ell}{\partial\nabla\varphi}-\vec{v}_s\frac{\partial\ell}{\partial\dot{\varphi}}\right]\right]=[[\ell]]\,, \tag{2.42}$$

for the flux related to φ at the interface S_n. From a physical viewpoint this is related to the production of the associated integral quantity. Therefore, (2.42) is referred to as the *production condition*.

There is an alternative way to derive the above production condition, (2.42), along the lines of distribution theory as sketched out in Sect. A.1.5 of Appendix A.1, showing that the generalized form of the formalism can be understood in terms of standard Lagrange formalism.

2.3 Resulting equations of motion and matching conditions

The above ideas are now applied to the discontinuous Lagrangian (2.32).

2.3.1 Equations of motion

In Appendix A.1.3 the Euler-Lagrange equations of (2.32) resulting from variation with respect to the elementary fields are calculated, based on which the corresponding equations of motion are derived. The equations of motion for the observable fields, (A.14), (A.19), are:

$$\nabla \cdot \vec{u} = 0\,, \tag{2.43}$$

$$\mathrm{D}_t \vec{u} = -\frac{\nabla p}{\varrho_0} + \nu \Delta \vec{u} - \nabla V + \vec{f}_{\mathrm{n.e.}}\,, \tag{2.44}$$

where $\vec{f}_{\mathrm{n.e.}}$ is used as an abbreviation for:

$$\vec{f}_{\mathrm{n.e.}} := -\frac{\nu}{\omega_0}\left\{ \mathrm{i} \ln \sqrt{\frac{\bar{\chi}}{\chi}} \left[\nabla \mathrm{tr}\underline{D}^2 + \{\mathrm{D}_t + \nabla \otimes \vec{u}\}\, \Delta \vec{u} \right] + \{\mathrm{D}_t + \nabla \otimes \vec{u}\} \left[2\underline{D} \mathrm{Im} \frac{\nabla \chi}{\chi} \right] \right\}. \tag{2.45}$$

Equation (2.43) shows that the continuity equation is unchanged, whereas the equations of motion (2.44) differ from the original NS equations since they contain additional forces $\vec{f}_{\mathrm{n.e.}}$. According to (2.45) the latter contain the complex field of thermal excitation, χ, and therefore the corresponding evolution equation, (A.18):

$$\mathrm{D}_t \chi + \mathrm{i}\omega_0 \chi = \frac{\nu}{2\bar{\chi}} \mathrm{tr}\underline{D}^2\,, \tag{2.46}$$

has to be considered additionally in order to complete the equation set. Note that the implicit velocity representation (A.17), obtained by variation with respect to $\vec{u}$, remains "Clebsch-like" but involves additional terms dependant on χ and $\vec{u}$.

Reconsidering the aforementioned hypothesis of Anthony [11] as well as Zuckerwar and Ash [288] in order to identify the additional forces occurring in the equations of motion as contributions due to a deviation from the thermodynamic local equilibrium, one would expect a limit case leading to the classical dynamics, as already discussed at the end of Subsec. 2.1.1. Indeed, according to (2.45) the extra forces are diminished when increasing the parameter ω_0 and the physical dimension of ω_0 is a reciprocal of time, suggesting

its interpretation as a relaxation rate. This interpretation is underpinned by considering the limit $\omega_0 \to \infty$, leading to vanishing of the extra forces, $\vec{f}_{\text{n.e.}} \to 0$ and therefore to a full reproduction of the NS equations by means of (2.44). In this equilibrium limit, the evolution equation (2.46) becomes meaningless, since the field of thermal excitation does not appear in the equations of motion any more.

It is noteworthy that the limit $\omega_0 \to \infty$ can be applied successfully to equations (2.44), reproducing the NS equations, but cannot be applied directly to the Lagrangian (2.32); most likely because the mechanical equations (2.43), (2.44) are decoupled from thermodynamics, whereas in the viscosity term of the Lagrangian (2.32) the occurring mechanical and thermodynamical degrees of freedom are strictly coupled.

For finite but sufficiently large values of ω_0, the additional forces $\vec{f}_{\text{n.e.}}$ due to thermodynamic non-equilibrium remain small, compared to viscous, external and pressure forces. According to (2.45) they consist of a factor $\mathrm{i}\ln\sqrt{\bar{\chi}/\chi}$ fluctuating between $-\pi$ and π, of quadratic terms with respect to velocity gradients and of third order derivatives of the velocity.

2.3.2 Matching conditions

As shown in Subsec. 2.2.2, variation with respect to the elementary fields, except for χ, induces matching conditions (2.40) at each interface. These are:

$$\delta\phi: \quad 0 = -\vec{n}\cdot[[\varrho_0(\vec{u}-\vec{v}_s)]]\,, \tag{2.47}$$

$$\delta\alpha: \quad 0 = 0\,, \tag{2.48}$$

$$\delta\beta: \quad 0 = -\vec{n}\cdot[[\varrho_0\alpha(\vec{u}-\vec{v}_s)]]\,, \tag{2.49}$$

$$\delta\vec{u}: \quad 0 = \varrho_0\frac{\nu}{\mathrm{i}\omega_0}\vec{n}\left[\left[\ln\sqrt{\frac{\bar{\chi}}{\chi}}\underline{D}\right]\right]\,. \tag{2.50}$$

According to the first condition, (2.47), the normal component of mass flux density has to be continuous, which physically corresponds to the conservation of the mass passing the interface. By inserting (2.47) into condition (2.49) it reduces to $[[\alpha]] = 0$, implying continuity of the Clebsch variable α. In order to understand the physics behind condition (2.50), it has to be taken into account that at each discontinuity the phase of the thermal excitation jumps from $\pm\pi$ to $\mp\pi$. In any case, the sign of $\mathrm{i}\ln\sqrt{\bar{\chi}/\chi}$ changes when turning from the rear side to the front side of an interface. As a consequence, condition (2.50) entails:

$$\vec{n}\,[\underline{D}]_- + \vec{n}\,[\underline{D}]_+ = \vec{0}\,, \tag{2.51}$$

which implies a reversal of the direction of the shear rate vector $\underline{D}\vec{n}$ along the discontinuous interface. Physically this is associated with slip occurring at the interface resulting from the absence of shear tangential to it, thus allowing fluid to migrate between neighbouring layers without friction as discussed in Sec. 2.4; the movement of the discontinuities giving rise to "slip waves". Once again this can be interpreted as a phenomenon due to a deviation from thermodynamic equilibrium.

2.3.3 Production condition and thermodynamic aspects

In order to apply formula (2.42) for the production condition, the complex field of thermal excitation has to be decomposed into modulus and phase according to $\chi = \sqrt{c_0 T}\exp(-\mathrm{i}\varphi)$, leading to the real-valued form:

$$\ell = -\varrho_0 \left[\mathrm{D}_t\phi + \alpha \mathrm{D}_t\beta - \frac{c_0 T}{\omega_0}\mathrm{D}_t\varphi - \frac{\vec{u}^{\,2}}{2} + c_0 T + V - \frac{\nu}{\omega_0} S(\varphi)\mathrm{tr}\underline{D}^2 \right] , \tag{2.52}$$

of the Lagrangian (2.32), where S again denotes the sawtooth function. The production condition then reads as:

$$\frac{\varrho_0}{\omega_0}\nabla\varphi \cdot [[c_0 T(\vec{u} - \vec{v}_s)]] = \varrho_0 \frac{\nu}{\omega_0}\left[\left[S(\varphi)\mathrm{tr}\underline{D}^2\right]\right] . \tag{2.53}$$

Despite the reversal of the shear rate tensor at the interfaces according to (2.51), its square, $\underline{D}^2$, remains continuous. Hence it follows that $\left[\left[S(\varphi)\mathrm{tr}\underline{D}^2\right]\right] = [[S(\varphi)]]\,\mathrm{tr}\underline{D}^2 = 2\pi\mathrm{tr}\underline{D}^2$, leading to:

$$\frac{1}{2\pi}\nabla\varphi \cdot [[c_0 T(\vec{u} - \vec{v}_s)]] = \nu\mathrm{tr}\underline{D}^2 . \tag{2.54}$$

The above condition reveals a discontinuity in the flux of the inner energy and therefore the production of inner energy due to dissipation at the interfaces. The latter can alternatively be related to the volume by estimating the gradient of the thermal phase as $\nabla\varphi \approx (2\pi/d)\vec{n}$, where d denotes the distance between two interfaces. As a consequence, the inhomogeneity at the right-hand side of the balance (2.54) can, according to $\nu\mathrm{tr}\underline{D}^2 \approx [[c_0 T(\vec{u} - \vec{v}_s)]]/d$, be re-interpreted as the mean production of inner energy related to the volume, at least in the sense of a statistical treatment.

Another source of inner energy production is given by equation (A.20), namely:

$$c_0 \mathrm{D}_t T = \nu\mathrm{tr}\underline{D}^2 , \tag{2.55}$$

which takes the form of a classical balance equation in continuum mechanics with a production rate $\nu\mathrm{tr}\underline{D}^2$ related directly to the volume. Compared with the classical theory of viscous flow [156, 241], the production of inner energy due to dissipation is twice the value

occurring in equation (2.55). Since, however, by (2.54) an additional production of inner energy at the inner boundaries is revealed, giving a contribution of the same amount as (2.55), the total production of inner energy is in accordance with classical theory.

The question arises as to whether the occurrence of discontinuous interfaces inside the fluid flow is an artefact of the model or if such phenomena really exist on a microscopic scale. Although a final answer to this question cannot be provided since knowledge about the processes occurring in a fluid flow on the micro-scale remains limited at the present time, it can be conjectured what kind of effect slight changes of the model may cause. At least the model established here accurately recovers the physics on a macroscopic scale, and there are dissipative phenomena with entropy production at discontinuous surfaces that have been known for a long time: with reference in particular to the classical theory of shock waves [241] where entropy production at discontinuous surfaces is provoked by a rapid compression of gas. Here, as explained earlier in the context of (2.51), the discontinuous phenomena take the form of slip waves.

Within this context it is also of particular interest how the physics would be affected by a change of branch cut for the complex logarithm: in Sec. 2.1.2 the standard branch cut along the negative real axis was used, leading to values of $\ln\sqrt{\bar{\chi}/\chi}$ between $-\mathrm{i}\pi$ and $\mathrm{i}\pi$. One consequence of this is that over time the fluctuating forces (2.45) occurring in the equations of motion statistically result in zero by averaging. If an alternative branch cut for the complex logarithm is considered, say e.g. a cut at arguments of the complex number if $a - \pi$, the positions of the discontinuous surfaces are shifted and the values of the complex logarithm go from $\mathrm{i}a - \mathrm{i}\pi$ to $\mathrm{i}a + \mathrm{i}\pi$. The first effect, the shift of the discontinuous surfaces, could be compensated for by applying the gauge transformation $\chi \to \chi \exp(-\mathrm{i}a)$ to the field of thermal excitation, whereas the second effect causes a change of the fluctuating forces (2.45) according to:

$$\vec{f}_{\mathrm{n.e.}} \longrightarrow \vec{f}_{\mathrm{n.e.}} + \frac{\nu a}{\omega_0}\left[\nabla \mathrm{tr}\underline{D}^2 + \{\mathrm{D}_t + \nabla \otimes \vec{u}\}\,\Delta\vec{u}\right] .$$

The extra term on the right side vanishes in the limit $\omega_0 \to \infty$; however, for finite values of ω_0 it provides an additional contribution to the equations of motion. By averaging again, this additional term does not tend to zero which is a valid physical argument for the use of the standard branch cut along the negative real axis (i.e. $a = 0$).

2.4 Discussion

It has been demonstrated that the dynamics resulting from the Lagrangian (2.32) can self-consistently be interpreted as an extension of the classical theory towards processes beyond thermodynamic local equilibrium. A reproduction of the classical theory is reached by

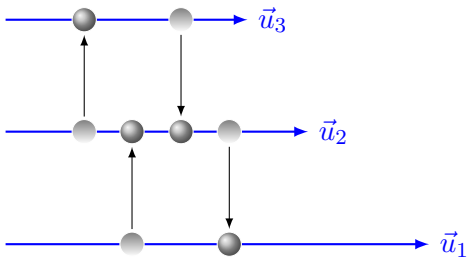


Figure 2.6: A simple microscopic model for viscosity, based on migration of particles between neighbored fluid layers by Brownian motion.

applying the limit $\omega_0 \to \infty$ for the relaxation rate to the equations of motion, a procedure that was not possible for the earlier suggested Lagrangian (2.4). In the following further indications are given in order to confirm the non-equilibrium assumption.

In the context of the Lagrangian (2.4), a striking non-classical feature is the difference between the momentum density and the mass flux density, namely the quasi momentum density $\vec{p}^{\,*} = \vec{p} - \varrho\vec{u}$. In the context of the Lagrangian (2.32), the mass flux density is given via (A.17), whereas the momentum density is obtained as a canonical Noether observable [218], giving: $\vec{p} = \varrho_0 \left[\nabla\phi + \alpha\nabla\beta + \frac{1}{\omega_0}\mathrm{Im}\left(\bar{\chi}\nabla\chi\right)\right]$. Hence, there is again a non-vanishing quasi momentum density,

$$\vec{p}^{\,*} = \varrho_0 \frac{\nu}{\omega_0} \nabla \cdot \left[\mathrm{i} \ln \sqrt{\frac{\bar{\chi}}{\chi}} 2\underline{D} \right] , \tag{2.56}$$

which in contrast to the quasi momentum density (2.7) resulting from the Lagrangian (2.4) tends to zero for the limiting case $\omega_0 \to \infty$. This is again in accordance with classical continuum theory.

Following the suggestion in [218, 222, 223] that the quasi-momentum takes into account contributions to the momentum beyond the scope of the continuum hypothesis on a molecular scale, e.g. Brownian motion, a possible physical interpretation of the quasi momentum density (2.7) is based on the elementary mechanism of viscosity, namely the transport of momentum by Brownian motion crosswise to the flow direction, see Fig. 2.6.

The viscosity of a fluid is usually explained on a molecular scale by an exchange of particles between neighboring fluid layers by Brownian motion of the molecules, by which a diffusion of momentum is induced. From the continuum viewpoint the migrating particles responsible for the diffusive momentum flux are “quasi-particles”, associated with an additional contribution to the momentum density. Hence, the quasi-momentum density can also be considered as a measure for the deviation from thermodynamic equilibrium.

Note that the interpretation of the additional terms in the equations of motion as physical non-equilibrium effects on a microscopic scale is also consistent with the weak

violation of the continuum hypothesis imposed by the discontinuous Lagrangian (2.32). Regardless, the violation of the continuum hypothesis vanishes in the limit $\omega_0 \to \infty$ for the relaxation rate: for increasing ω_0, the discontinuities are decreasing and for very large ω_0 they are physically reduced to fluctuations on a micro-scale in accordance with classical theory.

Finally, the reader is referred to the recent article by Marner et al. [173] comparing the above proposed variational principle with stochastic variational descriptions of dissipative systems, underpinning its physical basis from a different viewpoint. It is shown that additional non-classical contributions to the friction force occurring in the momentum balance vanish by time averaging. Accordingly, the discontinuous Lagrangian can alternatively be understood from the standpoint of an analogous deterministic model for irreversible processes of stochastic character. A comparison is made with established stochastic variational descriptions and an alternative deterministic approach based on a first integral of the NS equations is undertaken. The applicability of the discontinuous Lagrangian approach for different Reynolds number regimes is discussed considering the Kolmogorov time scale. A generalisation for compressible flow is elaborated and its use demonstrated for damped sound waves.

2.5 A generalized Clebsch transformation

Despite useful applications in fluid dynamics and other disciplines, the classical Clebsch transformation (1.23) has remained restricted to inviscid flow problems only. Although a digression from the main theme of this chapter, for completeness a generalisation of the transformation is developed allowing for the derivation of a generalised Bernoulli equation for viscous flow as published in Scholle and Marner [225].

Historically, the essential problem inhibiting the formulation of a Clebsch transformation for viscous flow is due to the friction force density $-\nu\Delta\vec{u}$ in the NS equations which does not, as mentioned in Chapter 1, fit into the scheme (1.28). Below the problem of finding a decomposition of the form (1.28) with prescribed Clebsch variables α, β is achieved for an *arbitrary* vector field $\vec{a}$ using a new approach. Consequently, a generalized Clebsch transformation is developed which encompasses the case of incompressible viscous flow. The resulting field equations are discussed briefly and solved for a representative flow problem.

2.5.1 Extension to viscous flow

(a) Non-applicability of classical Clebsch transformation to viscous flow

Assuming incompressible flow and constant kinematic viscosity the NS and continuity equations are:

$$\frac{\mathrm{D}\vec{u}}{\mathrm{D}t} - \nu\Delta\vec{u} + \nabla\left[\frac{p}{\varrho} + U\right] = \vec{0}\,, \tag{2.57a}$$

$$\nabla\cdot\vec{u} = 0\,, \tag{2.57b}$$

respectively [154]. In the more general case of compressible flow the continuity equation reads $\varrho\nabla\cdot\vec{u} = -\mathrm{D}\varrho/\mathrm{D}t$ and equations (2.57a) have to be replaced likewise by their more general form, frequently called the *Navier-Stokes-Duhem equations*, see e.g. [188].

The essential problem inhibiting the application of the Clebsch transformation to viscous flow is due to the friction force density $-\nu\Delta\vec{u}$ in the NS equations. Written in terms of the Clebsch variables, it reads:

$$-\nu\Delta\vec{u} = \nu\Delta\beta\nabla\alpha - \nu\Delta\alpha\nabla\beta - \nu\left(\nabla\alpha\cdot\nabla\right)\nabla\beta + \nu\left(\nabla\beta\cdot\nabla\right)\nabla\alpha\,. \tag{2.58}$$

Obviously, only two of the above four terms fit into the scheme (1.28), whereas the other two, subsumed to a vector field:

$$\vec{a} := \nu\left(\nabla\beta\cdot\nabla\right)\nabla\alpha - \nu\left(\nabla\alpha\cdot\nabla\right)\nabla\beta\,, \tag{2.59}$$

are of a mathematical form incompatible with (1.28). More generally, the problem of finding a decomposition of the form (1.28) with prescribed Clebsch variables α, β can be addressed via an *arbitrary* vector field $\vec{a}$ as shown below.

(b) Solution procedure

Introduce an auxiliary field ξ, fulfilling the following first order PDE:

$$\vec{\omega}\cdot\nabla\xi = \vec{\omega}\cdot\vec{a}\,, \tag{2.60}$$

with the vorticity ω given according to (1.32). This implies the identity:

$$\vec{\omega}\times\left(\vec{\omega}\times\left[\vec{a}-\nabla\xi\right]\right) = \vec{\omega}\left(\vec{\omega}\cdot\left[\vec{a}-\nabla\xi\right]\right) - \left[\vec{a}-\nabla\xi\right]\vec{\omega}^2 = -\vec{\omega}^2\left[\vec{a}-\nabla\xi\right]\,, \tag{2.61}$$

and therefore the decomposition of the difference $\vec{a}-\nabla\xi$ as:

$$\vec{a}-\nabla\xi = \frac{\left(\vec{\omega}\times\left[\vec{a}-\nabla\xi\right]\right)\times\vec{\omega}}{\vec{\omega}^2} = \frac{\left(\vec{\omega}\times\left[\vec{a}-\nabla\xi\right]\right)\times\left(\nabla\alpha\times\nabla\beta\right)}{2\vec{\omega}^2}\,,$$

$$= \frac{(\vec{\omega} \times [\vec{a} - \nabla\xi]) \cdot \nabla\beta}{2\vec{\omega}^2} \nabla\alpha - \frac{(\vec{\omega} \times [\vec{a} - \nabla\xi]) \cdot \nabla\alpha}{2\vec{\omega}^2} \nabla\beta \,; \tag{2.62}$$

i.e. as a linear combination of $\nabla\alpha$ and $\nabla\beta$. The decomposition (2.62) can be applied to any arbitrary vector field $\vec{a}$ in order to reach the form (1.28).

Like the Clebsch variables ϕ, α, β, the auxiliary field ξ is not given uniquely, since any particular solution ξ_p of the inhomogeneous linear first order PDE (2.60) can be superposed with any solution ξ_h of the respective homogeneous PDE $\vec{\omega} \cdot \nabla\xi_\mathrm{h} = 0$. Since three independent solutions are given by α, β and t, the mathematical theory of linear first order PDEs implies $\xi_\mathrm{h} = F(\alpha, \beta, t)$ for an arbitrary function F. As a consequence;

$$\xi \longrightarrow \xi' = \xi + F(\alpha, \beta, t) \,, \tag{2.63}$$

is a gauge transformation for the auxiliary field which is used subsequently giving a favourable form of the resulting equations.

(c) First integral of the Navier-Stokes equations

The NS equations (2.57a) contain identical mathematical terms as Euler's equations (1.27), apart from the pressure function taking the form $P = p/\varrho$ for incompressible flow, plus the friction term $-\nu\Delta\vec{u}$. Therefore, the Clebsch transformation leads to the three equations (1.29)-(1.31), supplemented with terms resulting from the decomposition of the friction term according to (2.58) and (2.60), (2.62). Using (2.63), the function F appearing in (1.29)-(1.31) is set to zero by gauging, leading finally to the set of the three scalar field equations:

$$\frac{\partial\phi}{\partial t} + \alpha\frac{\partial\beta}{\partial t} + \frac{\vec{u}^2}{2} + \frac{p}{\varrho} + U + \xi = 0 \,, \tag{2.64}$$

$$\frac{\mathrm{D}\alpha}{\mathrm{D}t} - \nu\Delta\alpha - \frac{\vec{\omega} \times [\vec{a} - \nabla\xi]}{2\vec{\omega}^2} \cdot \nabla\alpha = 0 \,, \tag{2.65}$$

$$\frac{\mathrm{D}\beta}{\mathrm{D}t} - \nu\Delta\beta - \frac{\vec{\omega} \times [\vec{a} - \nabla\xi]}{2\vec{\omega}^2} \cdot \nabla\beta = 0 \,, \tag{2.66}$$

where again $\vec{\omega}$ and $\vec{a}$ have been used as abbreviations according to (1.32), (2.59). Thus, a first integral of the NS equations has been constructed, based on the generalized Clebsch transformation: Eq. (2.64) is a generalization of Bernoulli's equation, whereas the two evolution equations (2.65), (2.66) for the vortex potentials α, β reveal the generic type of convection-diffusion equations with additional nonlinear coupling terms. The set of equations is completed by the PDE (2.60) for the auxiliary field ξ and the continuity equation (2.57b). The latter written in terms of Clebsch variables [195] is $\Delta\phi + \alpha\Delta\beta + 2\nabla\alpha \cdot \nabla\beta = 0$.

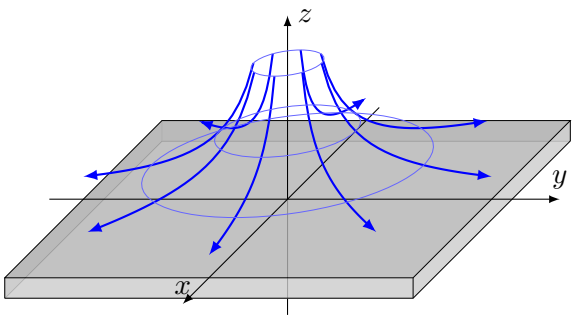


Figure 2.7: Schematic of an axisymmetric stagnation flow in the vicinity of a solid wall.

2.5.2 Axisymmetric stagnation flow as an example

Axisymmetric stagnation flow against a solid wall, see Fig. 2.7, is an example of prototypic character in fluid mechanics allowing for an analytical treatment of NS equations. In the inviscid case, the velocity field, written in cylindrical coordinates r, φ, z [175], reads:

$$\vec{u}_{\text{invis}} = ar\vec{e}_r - 2az\vec{e}_z \,. \tag{2.67}$$

Although (2.67) is a solution to the NS equations, it does not satisfy the no-slip condition $\vec{e}_r \cdot \vec{u} = 0$ at the wall and is therefore only a good approximation far from the wall. In the vicinity of the wall a boundary layer becomes manifest and the velocity is assumed to take a slightly different form, namely:

$$\vec{u} = rg(z)\vec{e}_r - 2f(z)\vec{e}_z \,,$$

with functions f, g that have to be determined. Associated boundary conditions are (i) the no-slip/no-penetration condition $\vec{u} = \vec{0}$ at $z = 0$ and (ii) the matching condition $\vec{u} \to \vec{u}_{\text{invis}}$ for $z \to \infty$, leading to:

$$f(0) = 0 \,, \qquad g(0) = 0 \,, \qquad \lim_{z\to\infty} f'(z) = a \,. \tag{2.68}$$

The field equations (2.57b), (2.60), (2.64)-(2.66) are solved as follows. First, the continuity equation (2.57b) implies: $g(z) = f'(z)$. Then, by reformulating the velocity according to

$$\vec{u} = f'(z)\nabla\frac{r^2}{2} - 2f(z)\nabla z = \nabla\left[\frac{r^2}{2}f'(z) - 2\int f(z)\mathrm{d}z\right] - \frac{r^2}{2}f''(z)\nabla z \,,$$

motivation is given for a Clebsch representation (1.23) with:

$$\phi = \frac{r^2}{2}f'(z) - 2\int f(z)\mathrm{d}z \,, \qquad \alpha = -\frac{r^2}{2}f''(z) \,, \qquad \beta = z \,. \tag{2.69}$$

Based on the latter the vector $\vec{a}$ is, according to (2.59), calculated as:

$$\vec{a} = \nu \frac{\partial}{\partial z} \nabla\alpha - 0 = \nu \nabla \frac{\partial \alpha}{\partial z} = -\frac{\nu}{2} \nabla \left[r^2 f'''(z) \right] .$$

The next step is determination of the auxiliary field ξ: as already discussed at the end of Sec. 2.5.1(b), the general solution of (2.60) is the sum of an arbitrary particular solution and the general solution, giving:

$$\xi = -\frac{\nu}{2} r^2 f'''(z) + F(\alpha, \beta, t) \,. \tag{2.70}$$

Hence, it follows that $\vec{a} - \nabla\xi = -\nabla F(\alpha, \beta, t)$ and, considering (1.32), that:

$$(\vec{\omega} \times [\vec{a} - \nabla\xi]) \cdot \nabla\beta = -\left(\left[\frac{\partial F}{\partial \alpha} \nabla\alpha + \frac{\partial F}{\partial \beta} \nabla\beta \right] \times \nabla\beta \right) \cdot \vec{\omega} = -2\vec{\omega}^2 \frac{\partial F}{\partial \alpha} \,,$$
$$(\vec{\omega} \times [\vec{a} - \nabla\xi]) \cdot \nabla\alpha = -\left(\left[\frac{\partial F}{\partial \alpha} \nabla\alpha + \frac{\partial F}{\partial \beta} \nabla\beta \right] \times \nabla\alpha \right) \cdot \vec{\omega} = 2\vec{\omega}^2 \frac{\partial F}{\partial \beta} \,.$$

Finally, the two field equations (2.65), (2.66) lead to:

$$2\nu f''(z) + \frac{r^2}{2} \left[\nu f''''(z) - 2f'(z) f''(z) + 2f(z) f'''(z) \right] - \frac{\partial F}{\partial \beta} = 0 \,, \tag{2.71}$$

$$-2f(z) + \frac{\partial F}{\partial \alpha} = 0 \,. \tag{2.72}$$

Since $z = \beta$, equation (2.72) can be integrated to give:

$$F(\alpha, \beta, t) = 2\alpha f(\beta) + F_1(\beta, t) \,,$$

with integration function F_1; which when inserted into (2.71), together with (2.69), leads to:

$$2\nu f''(z) - \frac{\partial F_1}{\partial \beta} + \frac{r^2}{2} \left[\nu f''''(z) + 2f(z) f'''(z) \right] = 0 \,.$$

By sorting terms with respect to powers of r, the above equation splits into two equations: $\partial F_1 / \partial \beta = 2\nu f''(z)$ and $\nu f''''(z) + 2f(z) f'''(z) = 0$. The latter equation can alternatively be written in integrable form as:

$$\frac{\mathrm{d}}{\mathrm{d}z} \left[\nu f'''(z) + 2f(z) f''(z) - f'(z)^2 \right] = 0 \,, \tag{2.73}$$

which following integration becomes: $\nu f'''(z) + 2f(z) f''(z) - f'(z)^2 = C$, where the integration constant C from the matching condition (2.68) is $C = -a^2$. Finally, via the substitution $f(z) = \sqrt{\nu a}\, \bar{f}(\bar{z})$ with $\bar{z} = \sqrt{a/\nu} z$, the whole problem is reduced to an ordi-

nary differential equation (ODE) of the form:

$$\bar{f}'''(\bar{z}) + 2\bar{f}(\bar{z})\bar{f}''(\bar{z}) - \bar{f}'(\bar{z})^2 + 1 = 0\,, \tag{2.74}$$

in accordance with the classical result. For more details and the solutions of (2.74) the reader is referred to [175]. Having solved the ODE (2.74), the generalized Bernoulli's equation (2.64) delivers the pressure p.

3 Approaches based on Goursat-like representation

In Chapter 1, Sec. 1.2.1(b), it is described how, by use of complex variables, a first integral of the 2D incompressible and steady NS equations can be established, the differential order of the former being lower. The procedure results in either a single complex valued equation of second order depending on a potential and the streamfunction (1.40a) or a system of two equations in the case when velocities are used; alternatively in terms of Cartesian coordinates a tensor formulation (1.41) can be found. While essentially a rediscovery of the result of Legendre [159], along similar lines to the work of Coleman [70] and Ranger [199], a hallmark of Scholle et al.'s [230] particular derivation is it provides a clear hint apropos generalisation to unsteady, 3D viscous flow: the attainment of which having hitherto remained out of reach, such a generalisation is the major focus of the present chapter.

3.1 Overview

It is important to appreciate the close relationship that exists between the above established first integral and the classical complex variable method (Goursat formulation) which is revealed when the streamfunction Ψ and the auxiliary potential field Φ are combined to form a complex field $\chi = \Phi + \mathrm{i}\eta\Psi$; in the Stokes flow case a simple bianalytic equation (1.45) for χ results, the solution of which is given by the combination $\chi = g_0(\xi) + \bar{\xi} g_1(\xi)$ of two holomorphic functions, as in (1.33). It is justifiable to take the first integral (1.44) to be a generalisation of the complex variable representation (1.45) towards viscous flows with inertia, with the difference that (1.44) no longer allows for a direct integration to the Goursat representation (1.33), Marner et al. [171].

Note, the potential field Φ is formally introduced as an auxiliary variable to make the field equations integrable, while posing the question as to its physical interpretation. Looking at the integrated equations in the Stokes flow case and the analogy between plane Stokes flow and plane linear elasticity based on a complex formulation, see [171], reveals that the first integral formulation reproduces the well-known complex Kolosov-Muskhelishvili formulas of linear elasticity, with the potential equivalent to an Airy stress

function except for a constant scaling; this is seen as follows. Using $\Phi = \mathrm{Re}(\chi)$ the representation of the potential derivatives in terms of the Goursat functions reproduce the Muskhelishvili-Kolosov formula [177, 184]:

$$\frac{\partial \Phi}{\partial x} + \mathrm{i}\frac{\partial \Phi}{\partial y} = w_0(\xi) + \xi\overline{w_1'(\xi)} + \overline{w_2'(\xi)}\,, \tag{3.1}$$

for the integrated components of stress in plane linear elasticity theory, revealing the potential to be closely connected with the Airy stress function.

At first glance, the analogy to linear elasticity theory provides an interpretation for the potential Φ in the linear Stokes case only, but rewriting the full NS equations (1.21a) in terms of the streamfunction and Airy stress function actually reproduces equation (1.41). Accordingly, the constitutive law of a Newtonian fluid involving the convective momentum flux density $R_{ij} = u_i u_j$[1] is adopted to account for inertial effects:

$$\sigma_{ij} = -p\delta_{ij} + \eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \varrho R_{ij}\ , \tag{3.2}$$

with σ_{ij} being a symmetric stress tensor, p the pressure and u_i the velocity components. Now introducing an Airy stress function such that:

$$\sigma_{11} = \frac{\partial^2 \Phi}{\partial y^2}\ , \quad \sigma_{22} = \frac{\partial^2 \Phi}{\partial x^2}\ , \quad \sigma_{12} = -\frac{\partial^2 \Phi}{\partial x \partial y}\ , \tag{3.3}$$

leads to:

$$-\varrho u_x u_y + \eta\left[\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right] + \frac{\partial^2 \Phi}{\partial x \partial y} = 0\ , \tag{3.4a}$$

$$\frac{\varrho}{2}\left[u_x^2 - u_y^2\right] + \eta\left[\frac{\partial u_y}{\partial y} - \frac{\partial u_x}{\partial x}\right] + \frac{1}{2}\left[\frac{\partial^2 \Phi}{\partial y^2} - \frac{\partial^2 \Phi}{\partial x^2}\right] = 0\ , \tag{3.4b}$$

which agrees with equation (1.41) except for a constant scaling of Φ. The above interpretation can be useful when considering more general flow problems.

In what follows, a first integral is constructed for the more general cases of unsteady and 3D flows. Firstly, in Sec. 3.2 it is shown that the first integral formulation of steady 2D-NS flow can be extended to transient flow if the real-valued potential field Φ is replaced by an appropriately defined complex potential, i.e. an extension of the formulation necessarily goes together with the introduction of new auxiliary fields. As a verification, the formulation is demonstrated to reproduce well-known velocity profiles of some standard test configurations – i.e. Couette flow, vortex decay and stagnation point flow. The steady case of Sec. 1.2.1 is recovered as a special case. It is shown how commonly en-

[1] By averaging, the Reynolds stress tensor is obtained from R_{ij}.

countered boundary conditions can be formulated in an elegant and useful form, with special attention paid to prescribing those occurring at a free surface. Motivated by the above integration procedure a representation of the dynamic boundary condition as a pure Dirichlet-Neumann condition on the potential is derived being a key feature of the new formulation.

Second, beginning with the transformation of existing 2D theory from a complex formulation to a real-valued one, resulting in the required tensor form (Sec. 3.3), the key aspect leading to the determination of an unsteady 3D first integral is recognition that it can be derived using a potential formulation similar to that employed in the reduction of Maxwell's equations. Via the astute use of gauge freedoms (Sec. 3.4), a decrease in the number of equations and unknowns is achieved as well as their transformation to a known, more tractable, equation set in which the differential order of the non-linear terms is reduced. Although consideration is focused on specific gauging of the tensor potential, in order to ensure the equation set has a favourable structure; the theory itself is amenable to alternative development. Some of which offer the prospect of a promising continuation of the research field; for example, it is shown that the gauge freedoms can be utilised intelligently to establish a variational principle for steady viscous flow. The limit case of inviscid flow is also addressed. Since the equations are derived in their most general form, restrictions to special cases such as steady or Stokes flow follows naturally, leading to further simplifications.

Boundary conditions, physical and auxiliary, in the framework of the above are provided, with the condition essential to the investigation of 3D free-surface flow problems derived in the form of a first integral of the usual dynamic boundary condition [171, 230]. As a whole the approach followed together with the established first integral, represents an important step forward; demonstrated via the solution of three classical, yet diverse, fluid flow problems of differing complexity, two of which are approached analytically, the other numerically (Sec. 3.5). In all three cases it is found that starting from the first integral, in deference to the NS equations, corresponding established solutions appearing in the open literature are recovered exactly; in one case it provides new theoretical insight. Last but not least, the time-evolution of periodically constrained unsteady flow is addressed as a standard scenario often encountered in relation to the direct numerical simulation (DNS) of viscous flow; using Fourier decomposition, the first integral formulation proves to be a very elegant approach leading to a reduced set of ODEs.

3.2 Two-dimensional unsteady and incompressible flow

The bulk of the work contained in this section stems from the following articles: Marner et al. [171] and Marner et al. [172].

3.2.1 Formulation of the field equations

Employing the complex variables ξ, $\overline{\xi}$ and the complex velocity field u, together with the streamfunction via (1.43), the unsteady NS equations (1.21a) can be transformed into the following scalar complex PDE:

$$\frac{\partial}{\partial\overline{\xi}}\left[-\mathrm{i}\varrho\frac{\partial\Psi}{\partial t}+p+\varrho\frac{\bar{u}u}{2}+U\right]+\varrho\frac{\partial}{\partial\xi}\left(\frac{u^2}{2}\right)=2\eta\frac{\partial^2 u}{\partial\overline{\xi}\partial\xi}\ ; \tag{3.5}$$

the continuity equation (1.21b) is fulfilled identically by (1.43).

(a) First integral using a complex potential of first order

By the introduction of a new complex potential M according to:

$$-\mathrm{i}\varrho\frac{\partial\Psi}{\partial t}+p+\varrho\frac{\bar{u}u}{2}+U=2\frac{\partial M}{\partial\xi}\ , \tag{3.6}$$

an integrable form of Eq. (3.5) is obtained:

$$\frac{\partial}{\partial\xi}\left[2\frac{\partial M}{\partial\overline{\xi}}+\varrho\frac{u^2}{2}-2\eta\frac{\partial u}{\partial\overline{\xi}}\right]=0\ , \tag{3.7}$$

which, following integration with respect to ξ, gives:

$$2\frac{\partial M}{\partial\overline{\xi}}+\varrho\frac{u^2}{2}-2\eta\frac{\partial u}{\partial\overline{\xi}}=f(\overline{\xi})\ . \tag{3.8}$$

The function $f(\overline{\xi})$ on the right-hand side of equation (3.8) can conveniently be set to zero by re-gauging the potential M according to:

$$M\longrightarrow M+\frac{1}{2}F(\overline{\xi}),\quad F'(\overline{\xi})=f(\overline{\xi})\ , \tag{3.9}$$

since expression (3.6) allows M to be augmented by an arbitrary $\overline{\xi}$-dependent complex function. Utilising the streamfunction (1.43), equation (3.8) simplifies to:

$$2\frac{\partial}{\partial\overline{\xi}}\left[M+2\mathrm{i}\eta\frac{\partial\Psi}{\partial\overline{\xi}}\right]+\varrho\frac{u^2}{2}=0\,. \tag{3.10}$$

(b) Representation by a potential of second order

From a pure technical viewpoint, a first integral has been successfully constructed in the form (3.10), since following differentiation with respect to ξ the NS equations are recovered. A comparison with the first integral (1.44) for steady flow, however, does not reveal an obvious relationship between both equations. In order to illustrate the

relationship, another complex potential χ is introduced fulfilling:

$$M + 2\mathrm{i}\eta\frac{\partial\Psi}{\partial\bar{\xi}} = 2\frac{\partial\chi}{\partial\bar{\xi}}\,, \tag{3.11}$$

leading directly to the compact form of the first integral equation (3.10):

$$4\frac{\partial^2\chi}{\partial\bar{\xi}^2} + \varrho\frac{u^2}{2} = 0\,, \tag{3.12}$$

which matches the first integral (1.44) for steady flow perfectly. Surprisingly, there is no difference between the complex field equations whether the flow is steady or unsteady. Indeed, the unsteady character of the flow is captured by the PDE (3.6) for the potential M, which after considering (3.11) reads as:

$$-\,\mathrm{i}\left[\varrho\frac{\partial\Psi}{\partial t} - 4\eta\frac{\partial^2\Psi}{\partial\bar{\xi}\partial\xi}\right] + p + \varrho\frac{\bar{u}u}{2} + U = 4\frac{\partial^2\chi}{\partial\bar{\xi}\partial\xi}\ . \tag{3.13}$$

Via (3.12), (3.13), two complex equations for one complex potential χ, the real-valued streamfunction Ψ and the pressure p arise.

3.2.2 Formulation of boundary conditions

Consider a simply connected domain with a boundary $x_i = f_i(s,t)$, parametrized with respect to the arc length s of the boundary. Furthermore, normal and tangential unit vectors along the boundary appear as:

$$t_i(s,t) = f_i'(s,t)\,, \tag{3.14}$$

$$n_i(s,t) = \varepsilon_{ji}t_j(s,t)\,, \tag{3.15}$$

where ε_{ij} denotes the Levi-Civita symbol. In complex notation, the tangential vector is given as $f'(s,t)$ with $f(s,t) = f_1(s,t) + \mathrm{i}f_2(s,t)$ and $n = \mathrm{i}f'(s,t)$; note that $\bar{n}n = \bar{t}t = \bar{f}'f' = 1$.

(a) No-slip and no-penetration condition at solid walls

In terms of the streamfunction Ψ, the no-slip and no-penetration condition along a solid wall $\xi = f(s)$ at rest take the usual form:

$$\mathrm{Im}\left(f'\frac{\partial\Psi}{\partial\xi}\right) = 0\,, \tag{3.16}$$

$$\Psi = \mathrm{const}\,, \tag{3.17}$$

of Dirichlet/Neumann boundary conditions.

(b) Kinematic and dynamic boundary conditions at a free surface

In the case of a free surface $\xi = f(s,t)$ movement of the surface, given by $\dot{f}$, has to be considered. In general, the movement $\dot{f}$ and the flow velocity u at the surface are not identical, but their normal components are. Hence, their difference is strictly directed tangential to the surface, i.e. $u - \dot{f} = u_1 f'$ with $u_1 \in \mathbb{R}$, implying after multiplication with $\bar{f}'$ and taking the imaginary part that:

$$u\bar{f}' - \bar{u}f' = \dot{f}\bar{f}' - \dot{\bar{f}}f' \ . \tag{3.18}$$

Equation (3.18) constitutes a kinematic boundary condition taking into account the coupling between the movement of the free surface and the flow velocity in the normal direction. Making use of relationship (1.43), this results in the following relationship:

$$\frac{\partial \Psi}{\partial s} = \frac{\mathrm{i}}{2}\left[\dot{f}\bar{f}' - \dot{\bar{f}}f'\right] =: \frac{\mathrm{i}}{2}\left\{f, \bar{f}\right\} \ , \tag{3.19}$$

where $\{\cdot,\cdot\}$ denotes the Poisson bracket of two functions [108].

The dynamic boundary condition in its original form reads:

$$\left[(p_0 - p)\delta_{ij} + \eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)\right]_{x_k = f_k(s)} n_j = \sigma \frac{\partial t_i}{\partial s} \ , \tag{3.20}$$

with σ on the right-hand side denoting the surface tension and p_0 the ambient pressure. Transformation of (3.20) into a complex representation provides the condition:

$$(p_0 - p)n - 2\eta \mathrm{i}\frac{\partial u}{\partial \bar{\xi}}\bar{f}' = \sigma f'' \ , \tag{3.21}$$

at $\xi = f(s,t)$. Next, by multiplying equation (3.6) with n, equation (3.10) with $\bar{n}$ and evaluating them at $\xi = f(s,t)$, their sum together with (3.21), after some manipulation, gives:

$$\varrho \frac{\partial \Psi}{\partial t} f' - \varrho u \frac{\partial \Psi}{\partial s} + Un = \frac{\partial}{\partial s}\left[\sigma f' + 2\mathrm{i}M - \mathrm{i}p_0 f\right] - \sigma' f' \, , \tag{3.22}$$

as the boundary condition for the complex potential M. The term $\mathrm{i}p_0 f$ can be eliminated by re-gauging the potential energy according to $U \to U - p_0$. By considering the kinematic boundary condition (3.19), the first two terms in equation (3.22) can be re-arranged as follows:

$$\begin{aligned} \frac{\partial \Psi}{\partial t} f' - u\frac{\partial \Psi}{\partial s} &= \frac{\partial \Psi}{\partial t} f' - \frac{\partial \Psi}{\partial \bar{\xi}}\left\{f, \bar{f}\right\} = \left[\frac{\partial \Psi}{\partial t} + \frac{\partial \Psi}{\partial \xi} f' + \frac{\partial \Psi}{\partial \bar{\xi}}\dot{\bar{f}}\right] f' - \frac{\partial \Psi}{\partial s}\dot{f} \\ &= \frac{\partial}{\partial t}\Psi\left(f, \bar{f}, t\right)\frac{\partial f}{\partial s} - \frac{\partial}{\partial s}\Psi\left(f, \bar{f}, t\right)\frac{\partial f}{\partial t} = \left\{\Psi(f, \bar{f}, t), f\right\} \ . \end{aligned}$$

Finally, by choosing the representation (3.11) in terms of the complex potential χ, the

dynamic boundary condition (3.22) can be re-written elegantly as:

$$\varrho \left\{ \Psi(f, \bar{f}, t), f \right\} + Un = \frac{\partial}{\partial s} \left[\sigma f' + 4\mathrm{i} \frac{\partial}{\partial \bar{\xi}} \left(\chi - \mathrm{i}\eta\Psi \right) \right] - \sigma' f' \,. \tag{3.23}$$

For flow conditions, for which the Poisson bracket in (3.23) vanishes (e.g. in the case of a steady flow) and the external force is time-independent, equation (3.23) becomes integrable; the associated first integral reads:

$$\sigma f' + 4\mathrm{i} \frac{\partial}{\partial \bar{\xi}} \left(\chi - \mathrm{i}\eta\Psi \right) = \int_{s_0}^{s} \left[U(\tilde{s}) n(\tilde{s}) + \sigma'(\tilde{s}) f'(\tilde{s}) \right] \mathrm{d}\tilde{s} \,. \tag{3.24}$$

The real-valued form of (3.24) reads:

$$2\varepsilon_{ij} \partial_j \Phi(s) = \sigma(s) t_i(s) - \int_{s_0}^{s} \left[U(\tilde{s}) n_i(\tilde{s}) + \sigma'(\tilde{s}) t_i(\tilde{s}) \right] \mathrm{d}\tilde{s} \,, \tag{3.25}$$

where the ε-symbol denotes $\varepsilon_{11} = \varepsilon_{22} = 0$, $\varepsilon_{12} = 1$, $\varepsilon_{21} = -1$, and reproduces the result of [230]. Note that in (3.25) the dynamic boundary condition, which is essentially a condition on the stresses, takes the form of a pure gradient condition on the potential field Φ which is not surprising when the close relationship between Φ and an Airy stress function, Sec. 3.1, is considered.

Knowledge of the potential gradient at the free surface allows the derivation of Dirichlet and Neumann boundary conditions for the potential variable as shown in [171]. By taking the inner product of (3.25) with t_i, its tangential component takes the form:

$$2 n_i(s) \partial_i \Phi(s) = \sigma(s) - t_i(s) \int_{s_0}^{s} \left[U(\tilde{s}) n_i(\tilde{s}) + \sigma'(\tilde{s}) t_i(\tilde{s}) \right] \mathrm{d}\tilde{s} \,, \tag{3.26}$$

which is a Neumann boundary condition for the potential Φ. Alternatively, the inner product of (3.25) with n_i leads to the corresponding normal component:

$$2 t_i(s) \partial_i \Phi(s) = n_i(s) \int_{s_0}^{s} \left[U(\tilde{s}) n_i(\tilde{s}) + \sigma'(\tilde{s}) t_i(\tilde{s}) \right] \mathrm{d}\tilde{s} \,. \tag{3.27}$$

Thus, using (3.27), (3.14)-(3.15) and partial integration, an integrated form of a Dirichlet boundary condition for the potential Φ can be constructed using the abbreviation $\gamma_i(s) = U(s) n_i(s) + \sigma'(s) t_i(s)$:

$$\begin{aligned}
\Phi(s) &= \Phi(s_0) - \frac{1}{2} \int_{s_0}^{s} \varepsilon_{ij} f_j'(\hat{s}) \int_{s_0}^{\hat{s}} \gamma_i(\tilde{s}) \, \mathrm{d}\tilde{s} \, \mathrm{d}\hat{s} \,, \\
&= \Phi(s_0) - \frac{1}{2} \left[\varepsilon_{ij} f_j(\hat{s}) \int_{s_0}^{\hat{s}} \gamma(\tilde{s}) \, \mathrm{d}\tilde{s} \right]_{\hat{s}=s_0}^{\hat{s}=s} + \frac{1}{2} \int_{s_0}^{s} \varepsilon_{ij} f_j(\tilde{s}) \gamma_i(\tilde{s}) \, \mathrm{d}\tilde{s} \,, \\
&= \Phi(s_0) + \frac{1}{2} \int_{s_0}^{s} \left[U(\tilde{s}) t_j(\tilde{s}) + \sigma'(\tilde{s}) t_i(\tilde{s}) \varepsilon_{ij} \right] \left[f_j(s) - f_j(\tilde{s}) \right] \mathrm{d}\tilde{s} \,.
\end{aligned} \tag{3.28}$$

By gauging, the term $\Phi(s_0)$ can be set to zero without loss of generality and moreover, in the case of $\sigma'(s) = 0$, that is neglecting Marangoni effects, further simplification is possible.

Derivation of the field equations (3.12), (3.13) in combination with boundary conditions of the form described in Sec. 3.2.2, represents a consistent extension of prior work concerning the integration of the incompressible NS equations [171, 230]. Here, the reduction of the dynamic boundary condition from its original form (3.20) to a standard Dirichlet-Neumann form (3.26), (3.28) is a key feature of the reformulation of the equations of motion in terms of the first integral of the NS equations allowing for the construction of efficient methods of analytical and numerical solution later on. In order to demonstrate the capabilities of this new approach and to focus on the distinguishing key feature, namely its ability to handle challenging transient flows, the non-trivial problem of Couette flow generated within an irregular flow geometry is considered in Sec. 6.4.

3.2.3 Application of the method to simple flows

In order to familiarise the reader with the methodology two well-known flow problems are considered, the solutions of which agree exactly with the standard ones existing in the literature.

(a) Uni-axial flow

For the simple case of a 2D uni-axial flow, $u_y = 0$, or equivalently in complex formulation $u - \bar{u} = 0$, the following PDE, cf. the streamfunction definition (1.43), has to be fulfilled:

$$\frac{\partial \Psi}{\partial \xi} + \frac{\partial \Psi}{\partial \overline{\xi}} = 0\,,$$

implying the explicit form:

$$\Psi = \Psi\left(\frac{\xi + \overline{\xi}}{2\mathrm{i}}, t\right) = \Psi\left(y, t\right)\,,$$
$$u = -2\mathrm{i}\frac{\partial \Psi}{\partial \overline{\xi}} = \Psi'\left(y, t\right)\,,$$

for the streamfunction and the velocity, respectively; where the prime denotes differentiation with respect to y. Next, the general solution of (3.12) reads:

$$\chi = \overline{\xi} w_0(\xi, t) + w_1(\xi, t) + \frac{\varrho}{2}\int\left[\int \Psi'\left(y, t\right)^2 \mathrm{d}y\right]\mathrm{d}y\,, \tag{3.29}$$

with two analytic functions $w_0(\xi, t)$, $w_1(\xi, t)$ being a generalization of the so-called *Goursat functions* [171]. By inserting (3.29) into equation (3.13) results in the following

simplification:

$$-\mathrm{i}\left[\varrho\dot{\Psi}-\eta\Psi''\right]+p+U=4w_0'(\xi,t)\,, \tag{3.30}$$

where a dot over a symbol denotes its time-derivative.

As an example involving the use of the above equation consider the problem of a horizontal plate of infinite extent covered by a fluid with flow invoked by an oscillatory movement, see Fig. 3.1. The no-penetration condition $\Psi(0,t)=0$ and the no-slip condition $\Psi'(0,t)=U\cos(\omega t)$ close the problem while in addition the asymptotic condition $\Psi'(y,t)\to 0$ has to be fulfilled. Assuming vanishing pressure $p=0$ and a wave-like solution of the form:

$$\Psi=U\,\mathrm{Im}\left\{\exp(-\mathrm{i}\omega t)\frac{\exp(\mathrm{i}ky)-1}{k}\right\}\,,$$

satisfying the no-slip and no-penetration conditions, equation (3.30) leads to the identity:

$$-\mathrm{i}U\,\mathrm{Im}\left\{\frac{-\mathrm{i}\varrho\omega+\eta k^2}{k}\exp(\mathrm{i}[ky-\omega t])+\mathrm{i}\varrho\frac{\omega}{k}\exp(-\mathrm{i}\omega t)\right\}=4w_0'(\xi,t)\,,$$

which is fulfilled if k and ω satisfy the dispersion relation:

$$\mathrm{i}\omega=\frac{\eta}{\varrho}k^2\,,$$

for damped transversal waves and the Goursat function w_0 takes the form:

$$w_0=-\frac{\mathrm{i}}{4}\xi\eta U\,\mathrm{Im}\left[k\exp(-\mathrm{i}\omega t)\right]\,.$$

Hence, the classical result given in [241] is reproduced.

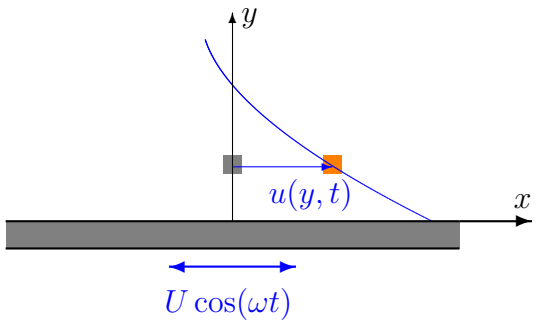


Figure 3.1: Geometry for the flow over an oszillating plate.

(b) Axisymmetric flow

Consider now a class of flows given by the following form:

$$\Psi=\Psi(r,t)\,, \tag{3.31}$$

$$u=-2\mathrm{i}\frac{\partial\Psi}{\partial\bar{\xi}}=-\frac{\mathrm{i}\xi}{r}\Psi'\,(r,t)\,, \tag{3.32}$$

$$r := \sqrt{\bar{\xi}\xi} \tag{3.33}$$

of the streamfunction; where the prime denotes derivation with respect to r, i.e. flows for which the streamlines of which are concentric circles, see Fig. 3.2a. Assuming a particular solution $\chi_\mathrm{p} = \chi_\mathrm{p}(r,t)$ of (3.12), leads to:

$$r\frac{\partial}{\partial r}\left(\frac{\chi_\mathrm{p}'}{r}\right) = \frac{\varrho}{2}\Psi'(r,t)^2 \,,$$

which, following integration, gives:

$$\frac{\chi_\mathrm{p}'}{r} = \frac{\varrho}{2}\int \frac{\Psi'(r,t)^2}{r}\mathrm{d}r \,.$$

By inserting this expression into equation (3.13), the following simplified PDE is obtained:

$$-\mathrm{i}\left[\varrho\dot{\Psi} - \frac{\eta}{r}\frac{\partial}{\partial r}\left(r\Psi'\right)\right] + p + U - \varrho\int \frac{\Psi'(r,t)^2}{r}\mathrm{d}r = 0 \,. \tag{3.34}$$

The introduction of the following similarity variable:

$$z = r\sqrt{\frac{\varrho}{\eta t}} \,,$$

enables one to search for solutions of the form $\Psi = f(z)$; in which case the imaginary part of equation (3.34) takes the form:

$$\frac{\varrho}{t}\left[\frac{z}{2}f'(z) + \frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\left(zf'(z)\right)\right] = 0 \,,$$

which after the substitution $g(z) = zf'(z)$ can be conveniently rewritten as:

$$g'(z) + \frac{z}{2}g(z) = 0 \,,$$

that has solutions of the form:

$$g(z) = g_0 \exp\left(-\frac{z^2}{4}\right) \,,$$

leading to:

$$f(z) = g_0\int \frac{1}{z}\exp\left(-\frac{z^2}{4}\right)\mathrm{d}z = \frac{g_0}{2}\,\mathrm{Ei}\left(-\frac{z^2}{4}\right) \,, \tag{3.35}$$

where Ei denotes the integral exponential function. This solution contains a singularity; however, by considering the superposition of (3.35) with the well-known potential vortex $\frac{\Gamma}{2\pi}\ln r$ as a different solution to (3.34), giving:

$$\Psi = \frac{g_0}{2}\,\mathrm{Ei}\left(-\frac{z^2}{4}\right) - \frac{\Gamma}{2\pi}\ln r \,,$$

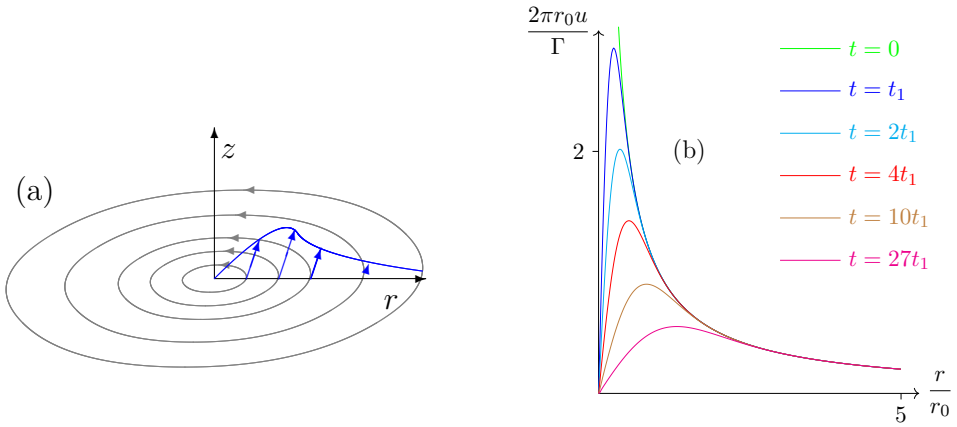


Figure 3.2: Vortex flow: (a) geometry; (b) time evolution of the velocity profile.

the singularity is removed as follows. The complex velocity field resulting from (3.32) reads:

$$u = -\frac{\mathrm{i}\xi}{r}\Psi'(r,t) = \frac{\mathrm{i}}{r}\left[\frac{\Gamma}{2\pi} + g_0 \exp\left(-\frac{\varrho r^2}{4\eta t}\right)\right]\exp(\mathrm{i}\varphi)$$

which is convergent in the limit $r \to 0$ if and only if $2\pi g_0 = \Gamma$. The final solution reads:

$$u = \frac{\mathrm{i}\Gamma}{2\pi r}\left[1 - \exp\left(-\frac{\varrho r^2}{4\eta t}\right)\right]\exp(\mathrm{i}\varphi)\ , \tag{3.36}$$

which is a reproduction of the classical Lamb-Oseen vortex [154]. The velocity profile $|u|$ is shown in Fig. 3.2b.

The above complex potential variant of the first integral does not lend itself readily to the solution of unsteady problems nor can it be extended to 3D problems. Accordingly, a different formulation is required which is the topic of the following section.

3.3 General three-dimensional unsteady and incompressible flow

With reference to the above work concerning the derivation and use of an exact complex-valued first integral for 2D incompressible flow [171, 172] a real-valued one for the full unsteady, incompressible NS equations is now formulated, as published in Scholle et al. [231]. Tensor calculus is employed, where vector fields are denoted by their Cartesian components, e.g. the velocity field $\vec{u}$ by u_i, $i = 1, \cdots, 3$, and tensors such as that for stress $\underline{T}$ by T_{ij}. The Einstein summation convention is used throughout: ∂_i denotes a spatial derivative with respect to x_i, i.e. $\partial_i = \partial/\partial x_i$, and ∂_t the time derivative; δ_{ij} is the Kronecker delta function and ε_{ijk} the 3D Levi-Civita symbol.

The beneficial use of potential fields, synonymous with Maxwell's theory [133], underpins the present approach: in that important and essential insight is gained for a similar treatment of the NS equations.

3.3.1 Preliminaries and introduction of a streamfunction vector

Prior to deriving the 3D form of the first integral, consideration is given to equation (1.21b), which is fulfilled identically by introducing a vector potential Ψ_k for the velocity according to:

$$u_i = \varepsilon_{ijk}\partial_j\Psi_k \,, \tag{3.37}$$

known in the literature as a 3D generalisation of the 2D streamfunction [22] that can be gauged by an arbitrary gradient field, that is:

$$\Psi_k \longrightarrow \Psi_k + \partial_k\chi \,; \tag{3.38}$$

leading, according to (3.37), to the same velocity field u_i.

Within the present context, the above allows reformulation of the time derivative in equation (1.21a) as the divergence of a tensor field, namely: $\partial_t u_i = \varepsilon_{ijk}\partial_t\partial_j\Psi_k = \partial_j\left(\varepsilon_{ijk}\partial_t\Psi_k\right)$. In this way equations (1.21a) can be re-written as:

$$\partial_j\left[\varrho\varepsilon_{ijk}\partial_t\Psi_k + \varrho u_j u_i - T_{ji} + U\delta_{ji}\right] = 0 \,, \tag{3.39}$$

with the stress tensor given by:

$$T_{ij} = -p\delta_{ij} + \eta\left[\partial_j u_i + \partial_i u_j\right] \,. \tag{3.40}$$

3.3.2 First integral of the field equations

With reference to the above it is clear that the momentum balance (3.39)) is a PDE of the same type as equation (1.21b) but for a tensor rather than a vector field. Hence, by introducing the tensor M_{lj} as a new potential, in accordance with:

$$\varrho\varepsilon_{ijk}\partial_t\Psi_k + \varrho u_i u_j - T_{ij} + U\delta_{ij} = \varepsilon_{jlk}\partial_l M_{ki} \,, \tag{3.41}$$

equation (3.39) is fulfilled identically; the analogy with (3.37) being obvious. Since T_{ij} is a symmetric tensor, it is convenient to split the above equation into symmetric:

$$\varrho u_i u_j - T_{ij} + U\delta_{ij} = \frac{1}{2}\left[\varepsilon_{jlk}\partial_l M_{ki} + \varepsilon_{ilk}\partial_l M_{kj}\right] \,, \tag{3.42}$$

and skew-symmetric parts; the latter, by multiplying (3.41) with ε_{ijn}, is conveniently represented as a vector equation:

$$2\varrho\partial_t\Psi_n = \partial_n M_{ll} - \partial_l M_{nl}\,. \tag{3.43}$$

Though not immediately obvious, the above rudimentary form of the first integral corresponds to that of the 2D first integral in Sec. 3.2; a more conveniently recognisable form is arrived at via the following reformulation.

First, using the streamfunction (3.37) the stress tensor (3.40) can be written as:

$$T_{ij} = -p\delta_{ij} + \eta\left[\varepsilon_{ilk}\partial_l\left(\partial_j\Psi_k\right) + \varepsilon_{jlk}\partial_l\left(\partial_i\Psi_k\right)\right]\,,$$

and, hence, equation (3.42) as:

$$\varrho u_i u_j + (p+U)\delta_{ij} = \frac{1}{2}\left[\varepsilon_{jlk}\partial_l\left(M_{ki} + 2\eta\partial_i\Psi_k\right) + \varepsilon_{ilk}\partial_l\left(M_{kj} + 2\eta\partial_j\Psi_k\right)\right]\,.$$

The first order potential M_{ki} enters the equations in combination with terms of the form $2\eta\partial_i\Psi_k$ only. Recognising this and following the procedure adopted earlier for 2D flow the combination $M_{ki} + 2\eta\partial_i\Psi_k$ can be rewritten as:

$$M_{ki} + 2\eta\partial_i\Psi_k = \varepsilon_{ipq}\partial_p a_{kq} + 2\partial_i\varphi_k\,, \tag{3.44}$$

which for vector fields is the well-known Maxwell decomposition into a divergence-free and a curl-free part; the form of which is a generalisation towards tensors of second rank with vector and tensor potential φ_k and a_{kq}, respectively. Inserting (3.44) into equations (3.42) and (3.43), yields the following relationships:

$$\varrho u_i u_j + (p+U)\delta_{ij} = \frac{1}{2}\varepsilon_{ilk}\varepsilon_{jpq}\partial_l\partial_p\left(a_{kq} + a_{qk}\right) + \partial_i\left(\varepsilon_{jlk}\partial_l\varphi_k\right) + \partial_j\left(\varepsilon_{ilk}\partial_l\varphi_k\right)\,, \tag{3.45}$$

$$2\varrho\partial_t\Psi_n = \partial_n\partial_k\left[\varepsilon_{kqp}a_{pq} + 2\varphi_k - 2\eta\Psi_k\right] - \partial_k\partial_k\left[2\varphi_n - 2\eta\Psi_n\right]\,, \tag{3.46}$$

which can be simplified by making use of the gauge transformation (3.38). The latter has no effect on equation (3.45) but equation (3.46) becomes:

$$2\varrho\partial_t\Psi_n = \partial_n\left\{\partial_k\left[\varepsilon_{kqp}a_{pq} + 2\varphi_k - 2\eta\Psi_k\right] - 2\varrho\partial_t\chi\right\} - \partial_k\partial_k\left[2\varphi_n - 2\eta\Psi_n\right]\,. \tag{3.47}$$

Since the gauge field χ can be chosen arbitrarily, the term $\varepsilon_{kqp}\partial_k a_{pq} - 2\varrho\partial_t\chi$ may be set to any value; in particular, by choosing:

$$\chi = \frac{1}{2\varrho}\int \varepsilon_{kqp}\partial_k a_{pq}\mathrm{d}t + \chi_0(x_i)\,, \tag{3.48}$$

leads to:

$$\varrho\partial_t\Psi_n = \partial_n\partial_k\left[\varphi_k - \eta\Psi_k\right] - \partial_k\partial_k\left[\varphi_n - \eta\Psi_n\right]\,; \tag{3.49}$$

showing that the skew-symmetric part of the tensor potential can be eliminated and therefore a_{pq} assumed symmetric from the very outset, leading ultimately to the following simplified form of equation (3.45):

$$\varrho u_i u_j + (p+U)\delta_{ij} = \varepsilon_{ilk}\varepsilon_{jpq}\partial_l\partial_p a_{kq} + \partial_i\left(\varepsilon_{jlk}\partial_l\varphi_k\right) + \partial_j\left(\varepsilon_{ilk}\partial_l\varphi_k\right) . \tag{3.50}$$

Second, the divergence $\partial_n(\cdots)$ of equation (3.49) leads to $2\varrho\partial_t\left(\partial_n\Psi_n\right) = 0$, implying that $\partial_n\Psi_n$ is independent of time. Since χ_0 in (3.48) is arbitrary, it can be chosen such that:

$$\partial_n\Psi_n = 0\,, \tag{3.51}$$

analogous to the Coulomb gauge in Maxwell's theory [133]. This, together with the identity $\partial_n\partial_k\varphi_k - \partial_k\partial_k\varphi_n = \varepsilon_{nij}\partial_i\left(\varepsilon_{jlk}\partial_l\varphi_k\right)$, enables equation (3.49) to be written in the form of an inhomogeneous diffusion equation:

$$\varrho\partial_t\Psi_n - \eta\partial_k\partial_k\Psi_n = \varepsilon_{nij}\partial_i\left(\varepsilon_{jlk}\partial_l\varphi_k\right)\,; \tag{3.52}$$

leading simultaneously to a reduction in the numbers of potentials due to the elimination of the skew-symmetric part of a_{pq}.

Thus far, a first integral of the unsteady incompressible Navier-Stokes equations has been obtained in the form of a tensor-valued field equation (3.50) and a vector-valued field equation (3.52) constrained by (3.51), involving various unknown fields a_{pq}, Ψ_n, u_n, p and φ_n. Although these remain to be closed mathematically, even at this stage they serve as an insightful starting point for fixing the remaining degrees of freedom in beneficial ways, that is *tuning* the form of the equations. This is explored in detail below.

3.4 Closure via selective gauge criteria

In general, a gauge transformation of a given set of potentials replaces them by an equivalent set of potentials leading to identical observables. Accordingly, such transformations can be used to simplify corresponding field equations, for the potentials, with respect to their mathematical structure as well as to the number of potentials. In the following, the gauge freedoms of a_{pq} and φ_n are analysed in detail. Obviously, by performing the operations:

$$a_{pq} \;\longrightarrow\; a_{pq} + \partial_p\alpha_q + \partial_q\alpha_p\,, \tag{3.53}$$

$$\varphi_n \;\longrightarrow\; \varphi_n + \partial_n\zeta\,, \tag{3.54}$$

for an arbitrary vector field α_q and an arbitrary scalar field ζ, the field equations (3.50), (3.52) remain invariant. The above rules are utilised subsequently to establish bona fide

gauging scenarios, ones that lead favourably to a reduction of the order of the established first integral, equations (3.50) to (3.52); in this context Scholle et al. [230] showed, by applying a particular gauge, that a special form of the first integral of NS equations for steady 3D flow can be obtained based on a minimum number of three potential fields only.

3.4.1 Convenient re-ordering of the first integral

Mixed derivatives of the form $\partial_k\partial_l(\cdots)$ are an inconvenience which can be avoided via a specific gauge transformation. This is achieved as follows, beginning with the re-ordering of the first the second order derivatives within the double curl operation $\varepsilon_{ikl}\varepsilon_{jpq}\partial_k\partial_p a_{lq}$ of equation (3.50). Since the product of two Levi-Civita symbols can be expressed as:

$$\varepsilon_{ikl}\varepsilon_{jpq} = \delta_{ij}\delta_{kp}\delta_{lq} + \delta_{ip}\delta_{kq}\delta_{lj} + \delta_{iq}\delta_{kj}\delta_{lp} - \delta_{ip}\delta_{kj}\delta_{lq} - \delta_{ij}\delta_{kq}\delta_{lp} - \delta_{iq}\delta_{kp}\delta_{lj}\,,$$

the identity:

$$\begin{aligned}\varepsilon_{ikl}\varepsilon_{jpq}\partial_k\partial_p a_{lq} &= -\partial_k\partial_k\left[a_{ij} - a_{ll}\delta_{ij}\right] + \partial_i\partial_k a_{kj} + \partial_j\partial_k a_{ki} - \partial_i\partial_j a_{kk} - \partial_l\partial_k a_{kl}\delta_{ij} \\ &= -\partial_k\partial_k\left[a_{ij} - \frac{a_{ll}}{2}\delta_{ij}\right] + \partial_i\partial_k\left[a_{kj} - \frac{a_{ll}}{2}\delta_{kj}\right] + \partial_j\partial_k\left[a_{ki} - \frac{a_{ll}}{2}\delta_{ki}\right] - \partial_l\partial_k\left[a_{kl} - \frac{a_{nn}}{2}\delta_{kl}\right]\delta_{ij}\,,\end{aligned}$$

results, giving rise to the following reformulation of equation (3.50):

$$\varrho u_i u_j + (p+U)\delta_{ij} = -\partial_k\partial_k\tilde{a}_{ij} + \partial_i A_j + \partial_j A_i - \partial_k A_k\delta_{ij}, \tag{3.55}$$

in terms of the modified tensor potential $\tilde{a}_{ij}$ and an auxiliary vector field A_j defined as:

$$\tilde{a}_{ij} := a_{ij} - \frac{a_{kk}}{2}\delta_{ij}\,, \tag{3.56}$$

$$A_j := \partial_k\tilde{a}_{kj} + \varepsilon_{jlk}\partial_l\varphi_k\,. \tag{3.57}$$

Note that from the form (3.55) of the tensor equation, the mathematical structure of the first integral for 2D flows demonstrated in Sec. 3.1 and 3.2 is recovered, see Appendix A.2.1.

Compared to its original form (3.50), equation (3.55) provides a partition of terms: in particular, all mathematical expressions with mixed derivatives of the form $\partial_k\partial_l(\cdots)$ occur exclusively as derivatives of the auxiliary vector field A_j. Now, via a gauge transformation of the form (3.53), the vector field A_j can be manipulated according to:

$$A_j \to A_j + \partial_k\partial_k\alpha_j\,, \tag{3.58}$$

which can be set to any arbitrary value by means of the proper choice of the gauge field α_j. The choice:

$$A_j = 0\,, \tag{3.59}$$

leads to the elimination of all mixed derivatives in (3.55), and to the simplified form:

$$\varrho u_i u_j + (p+U)\delta_{ij} = -\partial_k \partial_k \tilde{a}_{ij} \,. \tag{3.60}$$

The gauge condition (3.59) is reminiscent of the Lorenz gauge or Coulomb gauge in Maxwell's theory [133] which similarly leads to the elimination of mixed terms in the associated field equations. Moreover, via (3.59) the additional vector potential φ_j is eliminated from equation (3.60). By writing (3.59) explicitly as $\varepsilon_{jlk}\partial_l\varphi_k = -\partial_k\tilde{a}_{kj}$, φ_j can also be eliminated from equation (3.52), which accordingly takes the form:

$$\varrho\partial_t\Psi_n - \eta\partial_k\partial_k\Psi_n = -\varepsilon_{nkl}\partial_k\partial_m\tilde{a}_{ml} \,, \tag{3.61}$$

of an inhomogeneous diffusion equation; c.f. equation (3.52). Total elimination of φ_j from the entire set of equations requires the divergence of the gauge condition (3.59), implying:

$$\partial_j\partial_k\tilde{a}_{kj} = 0 \,. \tag{3.62}$$

The outcome is a favourably reduced equation set comprised of one each of a symmetric tensor equation (3.60), a vector equation (3.61) and a scalar equation (3.62), in terms of the symmetric modified tensor potential $\tilde{a}_{ij} = a_{ij} - a_{kk}\delta_{ij}/2$, the streamfunction vector Ψ_n and the pressure p.

The key features associated with equations (3.60)) to (3.62) are: (i) though the number of unknown fields (one symmetric tensor, one vector, one scalar) exceeds that of a comparable formulation in primitive variables (one vector and one scalar), their favourably different *structure* off-sets this; (ii) in contrast to the original NS equations (2.57a) which include the material time derivative, a non-linear term involving first order velocity derivatives, equation (3.60) consists of a non-linear term which depends directly on the velocities – equation (3.61) is simply a linear inhomogeneous diffusion equation, not a nonlinear diffusion-convection equation – resulting in a reduction of the differential order of the non-linearity.

(a) Zero-viscosity limit

Since the zero-viscosity limit leads to a change of problem type, namely from second order PDEs (NS equations) to ones of first order (Euler's equations), it is apposite to explore this special case: applying the limit $\eta \to 0$ to equations (3.60)-(3.62), the following set of PDEs:

$$\partial_k\partial_k\tilde{a}_{ij} = -\varrho u_i u_j - (p+U)\delta_{ij} \,, \tag{3.63}$$

$$\varrho\partial_t\Psi_n = -\varepsilon_{nkl}\partial_k\partial_m\tilde{a}_{ml} \,, \tag{3.64}$$

$$\partial_j\partial_k\tilde{a}_{kj} = 0 \,, \tag{3.65}$$

is obtained; still containing second order derivatives of the tensor potential but only first order derivatives of the streamfunction vector. Taking now the curl $\varepsilon_{pqn}\partial_q$ of (3.64), in combination with (3.63), (3.65) and (3.37), it follows that:

$$\varrho\partial_t \underbrace{(\varepsilon_{pqn}\partial_q\Psi_n)}_{u_p} = -\left[\delta_{pk}\delta_{ql} - \delta_{pl}\delta_{qk}\right]\partial_q\partial_k\partial_m\tilde{a}_{ml} = \partial_m\partial_k\partial_k\tilde{a}_{mp} - \partial_p\underbrace{\partial_l\partial_m\tilde{a}_{ml}}_{0}$$
$$= -\partial_m\left[\varrho u_m u_p + (p+U)\delta_{mp}\right] = -\varrho\overbrace{\partial_m u_m}^{0}u_p - \varrho u_m\partial_m u_p - \partial_p p - \partial_p U\,,$$

which is a full reproduction of Euler's equations, proving that the PDE set (3.63)-(3.65) is a first integral of Euler's equations, as it should be. Also in this case conservation of energy, momentum (in the absence of external forces, $U = 0$), angular momentum and helicity is fulfilled [22].

(b) Steady flow case

By employing the two gauge conditions (3.51), (3.62), together with the well known identity $\partial_k\partial_k\Psi_n = \partial_n\partial_k\Psi_k - \varepsilon_{nkl}\partial_k\left(\varepsilon_{lpq}\partial_p\Psi_q\right)$ and the definition of the streamfunction vector (3.37), equation (3.61) takes the form: $\varrho\partial_t\Psi_n + \varepsilon_{nkl}\partial_k\left[\eta u_l + \partial_m\tilde{a}_{ml}\right] = 0$. Hence, for steady flow, $\partial_t\Psi_n = 0$, the term in square brackets can be written as the gradient of a scalar field, that is:

$$\eta u_l + \partial_m\tilde{a}_{ml} = \partial_l\Phi\,;$$

by proper gauging of the tensor potential $\tilde{a}_{ml}$, Φ can be set equal to zero, resulting in the identity:

$$u_l = -\frac{1}{\eta}\partial_m\tilde{a}_{ml}\,, \tag{3.66}$$

via which the streamfunction vector is eliminated. The remaining fields are the symmetric tensor potential $\tilde{a}_{ml}$ and the pressure p; the field equations for steady flow being simply (3.60) and (3.62).

3.4.2 Traceless form

Two of the scalar fields, namely the pressure p and the trace of the tensor potential, can be eliminated as follows. The trace of equation (3.60):

$$\partial_k\partial_k\tilde{a}_{ii} = -\varrho u_i u_i - 3(p+U)\,, \tag{3.67}$$

enables direct calculation of the pressure from the other fields, c.f. Bernoulli's equation in potential theory. Equation (3.67) can be used to express $p + U$ in terms of the square of the velocity and second order derivatives of the tensor potential; this allows elimination of the pressure from equation (3.60), resulting in the following traceless symmetric tensor

equation:

$$\partial_k \partial_k \bar{a}_{ij} = -\varrho \left[u_i u_j - \frac{u_k u_k}{3} \delta_{ij} \right] , \tag{3.68}$$

in terms of the traceless tensor potential:

$$\bar{a}_{ij} = \tilde{a}_{ij} - \frac{\tilde{a}_{kk}}{3} \delta_{ij} . \tag{3.69}$$

Equation (3.68) is supplemented by equation (3.61) which in terms of the traceless tensor potential reads:

$$\varrho \partial_t \Psi_n - \eta \partial_k \partial_k \Psi_n = -\varepsilon_{nkl} \partial_k \partial_m \bar{a}_{ml} . \tag{3.70}$$

Together, expressions (3.68) and (3.70) comprise eight independent equations for the eight independent components of $\bar{a}_{ij}$ and Ψ_n, which is the minimum number in the case of unsteady flow.

Steady flow case: As above, a traceless and therefore reduced version of the field equations is achieved by inserting the identity (3.66) into (3.68) and taking (3.62) as the second equation, leading to six independent PDEs for six unknown fields.

3.4.3 Self-adjoint form

Finding variational formulations for physical systems is beneficial with respect to a deeper understanding of the system and for establishing new solution methods, both analytical and numerical. In fluid mechanics two major routes have emerged: (i) the stochastic variational description corresponding to the Lagrangian equations of motion in terms of material path lines, making use of a statistical treatment of kinetic models, see e.g. [13, 15, 63] and [14]; (ii) in the framework of a field description involving the recovery of the NS equations by variation of an action integral in the classical deterministic sense. In terms of the latter, it was Millikan [178] who showed the non-existence of a Lagrangian, in terms of the velocity u_i, the pressure p and their first order derivatives, that would enable the NS equations to be written as Euler-Lagrange equations. An analogue situation is found in Maxwell's theory, where it is not possible to establish a Lagrangian in terms of an electric field E_i and magnetic flux density B_i; however, a Lagrangian can be found in terms of a scalar potential φ and vector potential A_i. It is the latter that has prompted the search for a variational description for viscous flow in terms of potentials rather than velocity and pressure.

A variety of suggestions from different authors have appeared based on different potential formulations: Zuckerwar and Ash [288] used the Clebsch transformation [154, 191] to establish a Lagrangian for flows with volume viscosity, while latterly Scholle and Marner [226] consider shear viscosity in a similar manner. A variational description based on a

vector potential for the velocity was proposed by Bendali et al. [27]. In the present work the field equations are comprised of vector and tensor potentials, posing the question as to whether they are self-adjoint. As demonstrated below for the particular case of steady flow, a special gauge criteria is required to achieve a self-adjoint first integral of the NS equations.

When the flow is steady, $\partial_t \Psi_n = 0$, equation (3.49) is fulfilled identically by writing:

$$\varphi_n = \eta \Psi_n \,, \tag{3.71}$$

the insertion of which in equation (3.45) and making use of the relationship $a_{kq} + a_{qk} = 2\bar{a}_{kq} + 2a_{nn}\delta_{kq}/3$ from equation (3.69), leads to the following tensor equation:

$$\varrho u_i u_j + (p+U)\delta_{ij} = \varepsilon_{ilk}\varepsilon_{jpq}\partial_l\partial_p\bar{a}_{kq} + \frac{1}{3}\left[\partial_l\partial_l a_{nn}\delta_{ij} - \partial_i\partial_j a_{nn}\right] + \eta\left[\partial_i u_j + \partial_j u_i\right] \,, \tag{3.72}$$

as the most general form of the first integral for steady flow; valid for any gauging of the tensor potential. On elimination of the isotropic part and hence the pressure, its associated traceless form results:

$$\varrho\left[u_i u_j - \frac{u_k u_k}{3}\delta_{ij}\right] - \eta\left[\partial_i u_j + \partial_j u_i\right] = \left[\varepsilon_{ilk}\varepsilon_{jpq} - \varepsilon_{nlk}\varepsilon_{npq}\frac{\delta_{ij}}{3}\right]\partial_l\partial_p\bar{a}_{kq} + \frac{1}{3}\left[\partial_l\partial_l a_{nn}\delta_{ij} - \partial_i\partial_j a_{nn}\right] \,, \tag{3.73}$$

in terms of the traceless symmetric tensor potential $\bar{a}_{kq}$, the trace a_{nn} of the tensor potential and the velocity field $u_i = \varepsilon_{inm}\partial_n\Psi_m$. Suggesting a Lagrangian of the form:

$$\ell = \varrho\bar{a}_{ij}u_i u_j + \left[2\eta u_j - \frac{1}{3}\partial_j a_{nn}\right]\partial_i\bar{a}_{ij} + \frac{1}{2}\varepsilon_{ilk}\varepsilon_{jpq}\partial_l\bar{a}_{ij}\partial_p\bar{a}_{kq} + f\left(u_i, a_{nn}, \partial_i a_{nn}\right) \,, \tag{3.74}$$

which, because $u_i = \varepsilon_{inm}\partial_n\Psi_m$, is a function of the fields $\bar{a}_{kq}$, a_{nn}, Ψ_m and their associated first order derivatives, i.e. $\ell = \ell\left(\bar{a}_{kq}, a_{nn}, \partial_n\Psi_m, \partial_p\bar{a}_{kq}, \partial_i a_{nn}\right)$. f remains to be specified, its significance being discussed below.

Variation of the action integral:

$$\delta \iiint_V \ell\left(\bar{a}_{kq}, a_{nn}, \partial_n\Psi_m, \partial_p\bar{a}_{kq}, \partial_i a_{nn}\right) \mathrm{d}V = 0 \,, \tag{3.75}$$

with respect to $\bar{a}_{ij}$ results in the required Euler-Lagrange equations (3.73); whereas variation with respect to Ψ_m and a_{nn} lead to:

$$-2\varepsilon_{mni}\partial_n\left[\varrho\bar{a}_{ij}u_j + \eta\partial_j\bar{a}_{ji} + \frac{1}{2}\frac{\partial f}{\partial u_i}\right] = 0 \,, \tag{3.76}$$

$$\frac{1}{3}\partial_i\partial_j\bar{a}_{ij} + \frac{\partial f}{\partial a_{nn}} - \partial_i\left(\frac{\partial f}{\partial(\partial_i a_{nn})}\right) = 0 \,. \tag{3.77}$$

The meaning of equations (3.76) and (3.77) becomes much clearer after a substitution and rearrangement of terms: use of the definitions (3.56), (3.57) and (3.69) together with $\varepsilon_{jlk}\partial_l\varphi_k = \eta u_j$ following from (3.37) and (3.71), leads to the identity:

$$\partial_j \bar{a}_{ji} = A_i - \eta u_i + \frac{1}{6}\partial_i a_{nn}\,,$$

which, when substituted into (3.76), (3.77), yields:

$$\varepsilon_{mni}\partial_n A_i = \varepsilon_{mni}\partial_n \left[\eta u_i - \frac{1}{2\eta}\frac{\partial f}{\partial u_i} - \frac{\varrho}{\eta}\bar{a}_{ij}u_j\right], \tag{3.78}$$

$$\partial_i A_i = \partial_i \left[3\frac{\partial f}{\partial(\partial_i a_{nn})} - \frac{1}{2}\partial_i a_{nn}\right] - 3\frac{\partial f}{\partial a_{nn}}\,. \tag{3.79}$$

Since any vector field A_i can be reconstructed from its divergence $\partial_i A_i$ and its curl $\varepsilon_{mni}\partial_n A_i$, the reformulated Euler-Lagrange equations (3.78) and (3.79) is identifiable as an *alternative gauge* to that given by (3.59), the latter leading to the favourable formulation developed at the end of Section 3.4.1 for steady flow having a reduced number of unknown fields.

Hence, for steady flow a choice is available between the use of gauge (3.59) leading to a reduced set of fields and a favourable mathematical form of the field equations or gauge (3.76), (3.77) supplementing equations (3.73) to form a self-adjoint set of equations. The availability of a self-adjoint form can be useful for particular problems, e.g. when trying to compute normal forms around singular bifurcation points since it is necessary to make projections onto the eigenfunction of an adjoint problem, see Dijkstra et al. [85] and references therein.

Via a proper choice of the yet unknown function f in (3.74) the gauge conditions (3.78) and (3.79) are tuneable to some extent. For example, by choosing:

$$f\left(u_i, \partial_i a_{nn}\right) = \eta^2 u_i^2 + \frac{1}{12}\left(\partial_i a_{nn}\right)^2\,,$$

they simplify to:

$$\varepsilon_{mni}\partial_n A_i = -\frac{\varrho}{\eta}\varepsilon_{mni}\partial_n\left(\bar{a}_{ij}u_j\right)\,,$$
$$\partial_i A_i = 0\,.$$

In principle any arbitrary choice of f is possible. The variational principle above recovers, for $\eta \to 0$, the traceless version of the first integral of Euler's equations for steady flow.

3.5 Application of the Methodology

Having derived the first integral and explored its versatility in detail and on different levels, its use as a starting point to solve viscous fluid flow problems is now demonstrated. Not all of the gauge variants described in Sec. 3.4 are analysed further; rather the focus is those formulated in Sec. 3.4.1 and the solution of three different classical, benchmark viscous flows, which exhibit a hierarchy of sufficient complexity for such purposes – geometry, unsteadiness, non-linearity, inertia – and are solved analytically where analysis permits, otherwise numerically. The necessary, and related, boundary conditions required to do so are outlined below.

3.5.1 Boundary conditions

Depending on the problem of interest, the physical boundary conditions involved have to be formulated appropriately; a good example of this is the kinematic and dynamic boundary conditions required to solve 3D free surface flow problems which, although not utilised here, are included for completeness.

(a) Boundary conditions at solid walls, inlets and outlets

Along solid walls, for the velocity field the no-slip condition:

$$u_i = U_{Bi} \ , \tag{3.80}$$

has to be fulfilled, where U_{Bi} is the velocity of the boundary; inlet and outlet boundary conditions with fixed velocity profile have the same mathematical form as does the specification and advantageous use of symmetry and periodic boundary conditions. The latter type is discussed in more detail in Section 3.5.3(b).

(b) Boundary conditions at a free surface

Although a free surface condition does not appear in the problems solved below, the required attendant boundary conditions are provided. Their full derivation is given in Appendix A.2.2 but in summary two conditions must be fulfilled at a free surface: (i) the kinematic boundary condition, $u_i n_i = 0$, related to mass conservation; (ii) the dynamic boundary condition related to stress equilibrium at the surface. The latter can be described by the vector equation:

$$T_{ij} n_j = \sigma_s \kappa n_i \, , \tag{3.81}$$

involving the stress tensor T_{ij}, the normal vector n_i, the surface tension σ_s and the curvature κ. Using the potential representation for the respective physical quantities, equation

(3.81) can be reformulated into a more convenient form – see Appendix A.2.2, where it is also shown that for steady flow a first integral of the dynamic boundary itself can be constructed leading to a first order condition for the tensor potential entries only:

$$\varepsilon_{ikl}\left[\partial_k a_{lm}\mathrm{d}x_m + \left(\sigma_s n_k - \frac{U_k}{2}\right)\mathrm{d}x_l\right] = 0\,, \tag{3.82}$$

with the auxiliary functions U_k implicitly defined by (A.38).

(c) Auxiliary boundary conditions

Irrespective of the physical boundary conditions present, e.g. walls or free surfaces, an insufficient number can be prescribed to ensure a uniquely solvable system. An example of this is flow problems in which wall boundary conditions are prescribed on *all* parts of the boundary, as in the case of the lid-driven cavity flow explored below. Exactly three velocity conditions exist, which is less than the number of unknown fields. Even in the case of steady flow, where according to (3.66) the velocity can be expressed via the divergence of the tensor potential, 6 independent fields have to be considered – with at least three additional boundary conditions having to be formulated although there are no more physical conditions to be fulfilled; these necessary additional boundary conditions are subsequently termed *auxiliary boundary conditions* since they exert no influence on the physics.

While the options available for specifying these auxiliary boundary conditions appears wide, the two provided below are the only possible auxiliary Dirichlet conditions which appear reasonable:

1. Let n_j be the normal vector of the respective boundary. Then, three Dirichtlet boundary conditions are given by:

$$\tilde{a}_{ij}n_j = 0\,. \tag{3.83}$$

2. Let $t_i^{(1)}$ and $t_i^{(2)}$ be two orthogonal tangential vectors at the boundary. Then, three independent Dirichlet boundary conditions are given by:

$$t_i^{(1)}\tilde{a}_{ij}t_j^{(1)} = 0\,, \tag{3.84}$$
$$t_i^{(1)}\tilde{a}_{ij}t_j^{(2)} = 0\,, \tag{3.85}$$
$$t_i^{(2)}\tilde{a}_{ij}t_j^{(2)} = 0\,. \tag{3.86}$$

The decisive criterion for the choice of auxiliary boundary conditions is that they must not contradict the physically prescribed boundary conditions. For example, consider boundary

conditions (3.83) for a steady flow; by integration over the entire boundary ∂V of the system's volume V and making use of Gauss's theorem, the following identities:

$$0 = \oiint_{\partial V} \tilde{a}_{ij} n_j \mathrm{d}S = \iiint_V \partial_j \tilde{a}_{ij} \mathrm{d}V = -\eta \iiint_V u_i \mathrm{d}V \tag{3.87}$$

are obtained, where relationship (3.66) has been utilised. Equation (3.87) implies the vanishing of the global momentum, which is clearly an inadmissible physical restriction.

While the above example demonstrates the choice of auxiliary boundary conditions to be neither arbitrary nor intuitive, heuristic considerations lead to conditions (3.84)-(3.86) which do not conflict with the physics; although no proof is given, the comparatively accurate numerical results obtained below for the lid-driven cavity problem suggest the postulated conditions to be both admissible *and* sufficient to mathematically close the boundary value problem, at least in the steady case.

3.5.2 Unsteady stagnation flow

Consider the unsteady non-axisymmetric stagnation flow, depicted in to Figure 3.3, as a prototype example embodying both inertia and time dependence[2]. It is assumed that:

$$\vec{u} = xf'(z,t)\vec{e}_x + yg'(z,t)\vec{e}_y - [f(z,t)+g(z,t)]\,\vec{e}_z = \nabla \times [yg(z,t)\vec{e}_x - xf(z,t)\vec{e}_y]\,,$$

where the prime denotes differentiation with respect to z. Accordingly, the continuity equation (1.21b) is fulfilled identically and in which case the velocity can be obtained from a streamfunction vector, according to equation (3.37), with:

$$\vec{\Psi} = yg(z,t)\vec{e}_x - xf(z,t)\vec{e}_y\,; \tag{3.88}$$

note too, that the streamfunction vector fulfils the Coulomb gauge (3.51). The traceless form of the first integral is utilised, equations (3.68) and (3.70); written in component form equation (3.68) reads:

$$-\varrho^{-1}\Delta\bar{a}_{11} = \frac{2}{3}x^2 f'^2 - \frac{1}{3}y^2 g'^2 - \frac{1}{3}(f+g)^2\,, \tag{3.89}$$

$$-\varrho^{-1}\Delta\bar{a}_{22} = \frac{2}{3}y^2 g'^2 - \frac{1}{3}x^2 f'^2 - \frac{1}{3}(f+g)^2\,, \tag{3.90}$$

$$-\varrho^{-1}\Delta\bar{a}_{12} = xyf'g'\,, \tag{3.91}$$

$$-\varrho^{-1}\Delta\bar{a}_{13} = -x(f+g)f'\,, \tag{3.92}$$

$$-\varrho^{-1}\Delta\bar{a}_{23} = -y(f+g)g'\,, \tag{3.93}$$

[2]The example is similar to that considered in Sec. 2.5.2 but in contrast to the latter exhibits increased complexity because *unsteady* and *non*-axisymmetric flow is assumed; also the solution approach is completely different here.

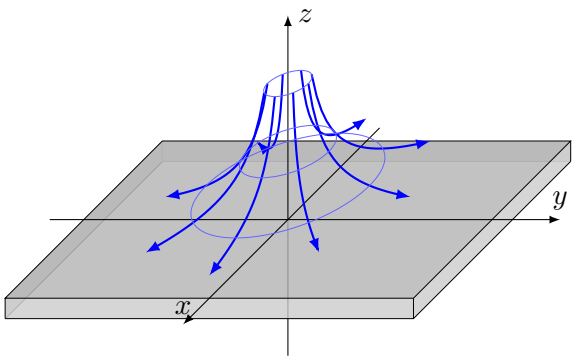


Figure 3.3: Schematic of a non-axisymmetric, unsteady stagnation flow problem close to a solid wall.

while equation (3.70) gives:

$$-y\varrho\left[\dot{g}-\nu g''\right]=\partial_1\partial_2\bar{a}_{13}-\partial_1\partial_3\bar{a}_{12}+\{\partial_2\partial_2-\partial_3\partial_3\}\bar{a}_{23}+\partial_2\partial_3[\bar{a}_{33}-\bar{a}_{22}]\,, \tag{3.94}$$

$$x\varrho\left[\dot{f}-\nu f''\right]=\partial_1\partial_3[\bar{a}_{11}-\bar{a}_{33}]+\partial_3\partial_2\bar{a}_{12}-\partial_1\partial_2\bar{a}_{23}+\{\partial_3\partial_3-\partial_1\partial_1\}\bar{a}_{13}\,, \tag{3.95}$$

$$0=\{\partial_1\partial_1-\partial_2\partial_2\}\bar{a}_{12}+\partial_1\partial_2[\bar{a}_{22}-\bar{a}_{11}]+\partial_1\partial_3\bar{a}_{23}-\partial_2\partial_3\bar{a}_{13}\,, \tag{3.96}$$

the dot above a symbol, here and subsequently, denoting differentiation with respect to time.

The boundary conditions at $z=0$ are the usual no-slip/no-penetration conditions $f'(0,t)=g'(0,t)=0$ and $f(0,t)=g(0,t)=0$. The characteristic nature of above viscous stagnation point flow is boundary layer like [175], with the requirement that it matches the associated potential flow as $z\to\infty$. Accordingly, the tensor potential for an inviscid boundary layer flow has to be constructed *a priori*.

(a) Associated potential flow

In the case of 3D stagnation flow, the corresponding potential flow is given [175] by: $f(z)=a_1z$, $g(z)=a_2z$, fulfilling the no-penetration condition $f(0)=g(0)=0$, but not so the no-slip condition. For the construction of the associated traceless tensor potential, equations (3.89)-(3.93) have to be solved. One such particular solution is given by:

$$\bar{a}^{\mathrm{p}}_{11}=-\frac{\varrho}{6}\left[2a_1^2x^2z^2-a_2^2y^2z^2-\frac{1}{6}(3a_1^2+2a_1a_2)z^4\right]\,, \tag{3.97}$$

$$\bar{a}^{\mathrm{p}}_{22}=-\frac{\varrho}{6}\left[2a_2^2y^2z^2-a_1^2x^2z^2-\frac{1}{6}(3a_2^2+2a_1a_2)z^4\right]\,, \tag{3.98}$$

$$\bar{a}^{\mathrm{p}}_{12}=-\frac{\varrho}{2}a_1a_2xyz^2\,,\quad \bar{a}^{\mathrm{p}}_{13}=\frac{\varrho}{6}a_1(a_1+a_2)xz^3\,,\quad \bar{a}^{\mathrm{p}}_{23}=\frac{\varrho}{6}a_2(a_1+a_2)yz^3\,; \tag{3.99}$$

which fulfils equation (3.96), but not equations (3.94), (3.95); a superposition of the form $\bar{a}_{ij} = \bar{a}^{\mathrm{h}}_{ij} + \bar{a}^{\mathrm{p}}_{ij}$ with $\Delta\bar{a}^{\mathrm{h}}_{ij} = 0$ is required in order to fulfil all of the equations. By choosing $\bar{a}^{\mathrm{h}}_{12} = \bar{a}^{\mathrm{h}}_{13} = \bar{a}^{\mathrm{h}}_{23} = 0$ and:

$$\bar{a}^{\mathrm{h}}_{11} = A_1 \left[x^4 + z^4 - 6x^2z^2\right] , \qquad \bar{a}^{\mathrm{h}}_{22} = A_2 \left[y^4 + z^4 - 6y^2z^2\right] ,$$

equations (3.94), (3.95) results in: $0 = [48A_2 + \varrho a_2^2]yz\,, \quad 0 = -[48A_1 + \varrho a_1^2]xz\,,$ implying $A_1 = -\varrho a_1^2/48$ and $A_2 = -\varrho a_2^2/48$; the other equations are not affected. Hence, the resulting solutions of the homogeneous equations read:

$$\bar{a}^{\mathrm{h}}_{11} = -\varrho\frac{a_1^2}{48}\left[x^4 + z^4 - 6y^2z^2\right] , \tag{3.100}$$

$$\bar{a}^{\mathrm{h}}_{22} = -\varrho\frac{a_2^2}{48}\left[y^4 + z^4 - 6y^2z^2\right] . \tag{3.101}$$

(b) General case

Assume the following analogous form of the traceless tensor potential:

$$\bar{a}^{\mathrm{p}}_{11} = -\varrho\left[F_{110}(z,t) + x^2F_{111}(z,t) + y^2F_{112}(z,t)\right] ,$$
$$\bar{a}^{\mathrm{p}}_{22} = -\varrho\left[F_{220}(z,t) + x^2F_{221}(z,t) + y^2F_{222}(z,t)\right] ,$$
$$\bar{a}^{\mathrm{p}}_{12} = -\varrho xyF_{12}(z,t)\,, \quad \bar{a}^{\mathrm{p}}_{13} = \varrho xF_{13}(z,t)\,, \quad \bar{a}^{\mathrm{p}}_{23} = \varrho yF_{23}(z,t)\,,$$

for the particular solution of (3.89)-(3.93); while remembering that as above the flow is of a boundary-layer type. In order to fulfil the matching condition, this particular solution has to be supplemented by equations (3.100), (3.101). In this way, equations (3.89)-(3.93) are reduced as follows:

$$3F''_{110} + 2F_{111} + 2F_{112} = -\frac{1}{3}(f+g)^2\,, \quad F''_{111} = \frac{2}{3}f'^2\,, \qquad F''_{112} = -\frac{1}{3}g'^2\,, \tag{3.102}$$

$$F''_{220} + 2F_{221} + 2F_{222} = -\frac{1}{3}(f+g)^2\,, \quad F''_{221} = -\frac{1}{3}f'^2\,, \qquad F''_{222} = \frac{2}{3}g'^2\,, \tag{3.103}$$

$$F''_{12} = f'g'\,, \qquad F''_{13} = (f+g)f'\,, \quad F''_{23} = (f+g)g'\,, \tag{3.104}$$

written in terms of functions F_{110}, F_{111}, F_{112}, F_{220}, F_{221}, F_{222}, F_{12}, F_{13}, F_{23}. By inserting the above solution into equations (3.94) to (3.96), it is found that equation (3.96) is fulfilled identically whereas equations (3.94) and (3.95) yield:

$$\dot{g} - \nu g'' = a_2^2 z + F''_{23} - F'_{12} - 4F'_{222} - 2F'_{112} = a_2^2 z + gg' + \int\left[fg'' - 2g'^2\right]\mathrm{d}z, \tag{3.105}$$

$$\dot{f} - \nu f'' = a_1^2 z + F''_{13} - F'_{12} - 4F'_{111} - 2F'_{221} = a_1^2 z + ff' + \int\left[gf'' - 2f'^2\right]\mathrm{d}z, \tag{3.106}$$

which, upon taking their derivative with respect to z, leads to a coupled set of third order equations for the functions $f(z,t)$ and $g(z,t)$, namely:

$$\dot{g}' - \nu g''' = a_2^2 - g'^2 + (f+g)g'' \,, \tag{3.107}$$

$$\dot{f}' - \nu f''' = a_1^2 - f'^2 + (f+g)f'' \,. \tag{3.108}$$

These have to be solved numerically; the special case of a steady flow, $\dot{f} = \dot{g} = 0$, results in a set of ODEs as reported and solved by Howarth [130].

3.5.3 Flow within a cubic domain

(a) Steady flow within a lid-driven cavity

The case of stationary viscous flow in a square-sided 3D lid-driven cavity of equal edge length, L, and a constant upper lid velocity of U_0 [86], is explored through the numerical solution of the primitive variable form of the first integral for steady flow tuned as per the corresponding gauge criterion of Section 3.4.1 – a key feature being that the essential equation (3.60) is devoid of mixed derivatives, with the consequent benefit it simplifies and accelerates the use of iterative solvers. The equations to be solved, namely (3.60), (3.62) and (3.66), when non-dimensionalised in terms of L and U_0, read:

$$\partial_k \partial_k \tilde{a}_{ij} + \mathrm{Re}\, u_i u_j + (p+U)\delta_{ij} = 0 \qquad \text{in } \Omega \,, \tag{3.109}$$

$$\partial_l \partial_k \tilde{a}_{kl} = 0 \qquad \text{in } \Omega \,, \tag{3.110}$$

$$-\partial_k \tilde{a}_{kl} = u_l \qquad \text{in } \overline{\Omega} \,, \tag{3.111}$$

where $\mathrm{Re} = \frac{\varrho U_0 L}{\eta}$ is the Reynolds number. $\overline{\Omega}$ in equation (3.111) denotes the closed set of the solution domain $\Omega = [0,1]^3$ with boundary $\partial\Omega$ (the moving lid lying in the plane $z = 1$) and indicates that (3.111) is valid both in the inner domain defining the velocities from the known tensor potential entries and at the boundary where the velocities are prescribed in the form of Dirichlet conditions, that is by $u_l = g_l$ on $\partial\Omega$ for appropriate g_l. Equations (3.109)-(3.111) are complemented by the three auxiliary Dirichlet boundary conditions (3.84)-(3.86) for the tensor potential entries in order to obtain a uniquely solvable equation set: although this remains to be proven formally, the numerical results indicate the above system to be mathematically closed.

Newton's method is employed to generate a sequence of $n \in \mathbb{N}_0$ linearised systems based on the following steps:

Step 1:

$$\partial_k \partial_k \tilde{a}_{ij}^{(n+1)} - \mathrm{Re}\left[u_i^{(n)} \partial_k \tilde{a}_{kj}^{(n+1)} + u_j^{(n)} \partial_k \tilde{a}_{ki}^{(n+1)}\right] + (p^{(n+1)} + U)\delta_{ij} = \mathrm{Re}\, u_i^{(n)} u_j^{(n)}\,, \quad (3.112)$$

$$\partial_l \partial_k \tilde{a}_{kl}^{(n+1)} = 0 \quad \text{in } \Omega\,, \quad (3.113)$$

$$-\partial_k \tilde{a}_{kl}^{(n+1)} = g_l \quad \text{on } \partial\Omega\,, \quad (3.114)$$

Step 2:

$$u_l^{(n+1)} := -\partial_k \tilde{a}_{kl}^{(n+1)} \qquad \text{in } \Omega\,, \quad (3.115)$$

in which (3.111) has been used to replace the velocities in (3.112) as primary unknowns with index $(n+1)$; the velocities $u_i^{(n)}$ in (3.112) are assumed known from the previous iterative step having been calculated from the tensor potential via (3.115). As such, the above equations only involve the six tensor potential entries and the pressure as independent primary variables, with the velocities appearing as secondary variables. Iteration starts from $n = 0$ where the unknown fields are initialised with respect to the linear Stokes flow solution.

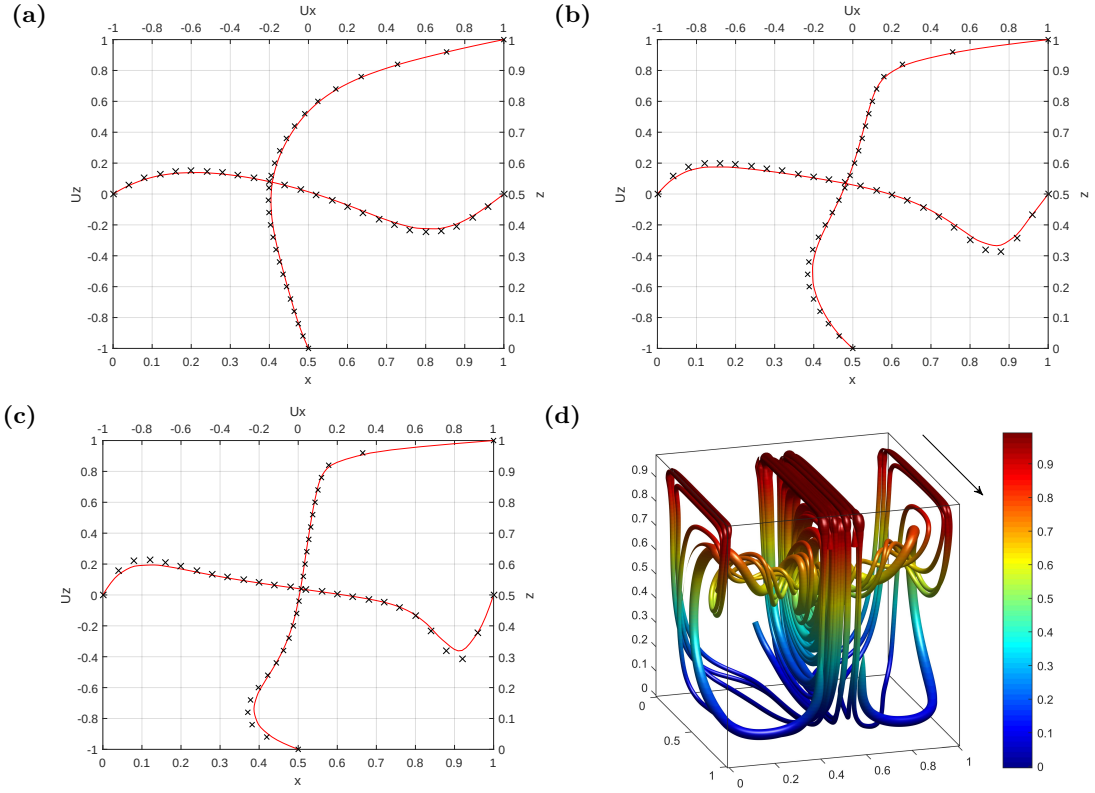


Figure 3.4: 3D lid-driven cavity flow. Centre-line velocity profiles for u_x and u_z in the plane intersections $x = y = 0.5$ and $y = z = 0.5$, respectively, for Reynolds numbers of (a) 100, (b) 400 and (c) 1000; the results from the present work are shown as solid red curves and compared to those of Ding et al. [86] shown as black crosses. (d) Shows selected stream tubes for the case Re = 400, the arrow indicating the direction of motion of the upper moving lid.

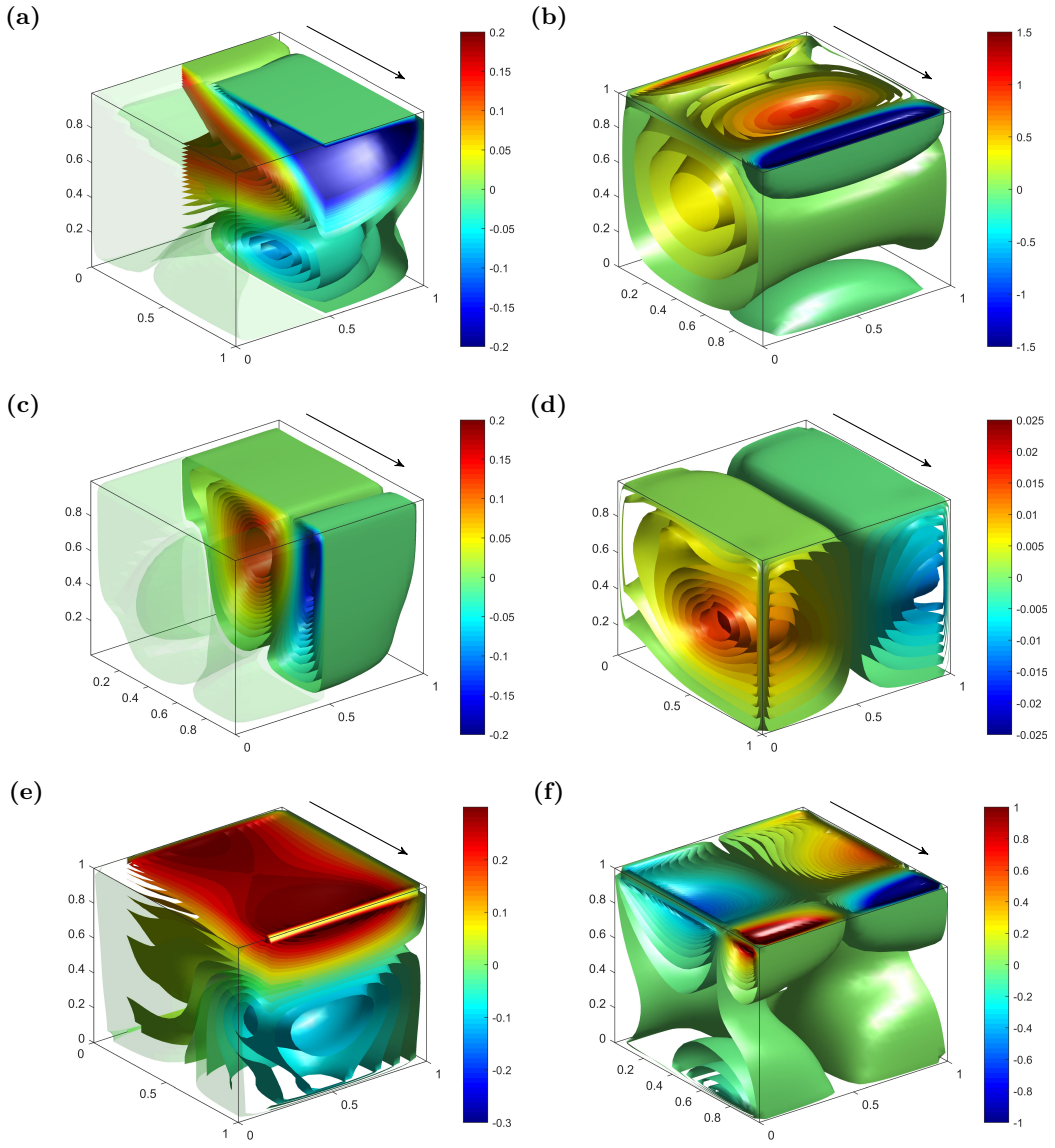


Figure 3.5: Visualisation of the six tensor potential entries for the 3D lid-driven cavity flow problem, for the case Re = 400: (a) a_{11}; (b) a_{22}; (c) a_{33}; (d) a_{12}; (e) a_{13}; (f) a_{23}. The arrow indicates the direction of motion of the upper moving lid.

For demonstration purposes the cubic nature of the flow field is well suited to solution via a finite difference methodology and structured Cartesian grid system which is the approach adopted. In doing so the well-known oscillatory pressure instability problem linked with the discretisation of flow problems in terms of primitive variables is avoided by employing a velocity-pressure staggered grid arrangement [95] which is extended to encompass the remaining unknowns, namely the tensor potential entries, in a consistent way. Accordingly, central difference stencils and therefore the discrete equations are well defined everywhere.

Although the numerical scheme itself is not the focus of the present work, as it is the first such implementation of the same in the present context the details are summarised in Appendix A.2.4. The above equations are similarly amenable to solution utilising, for example, a more complex irregular grid structure and finite element methodology that satisfies a compatibility condition between solution spaces when employing mixed finite elements [107]. Solutions are presented for three different Reynolds numbers up to and including Re = 1000.

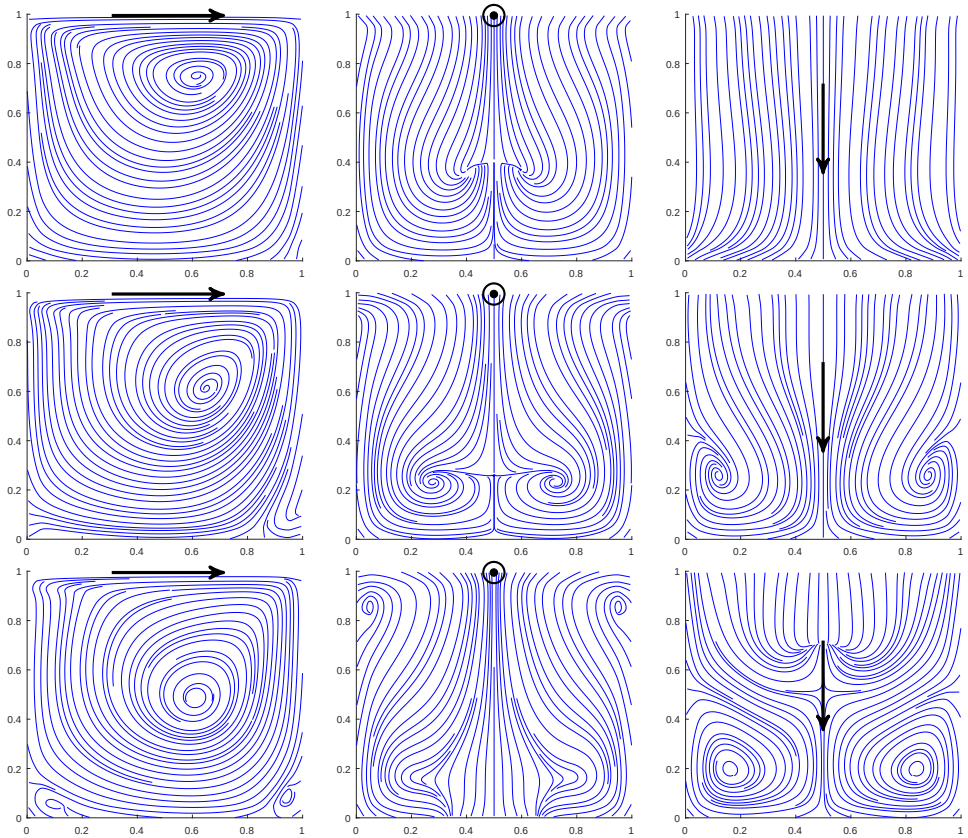


Figure 3.6: Streamline projections onto fixed planes for flow in a 3D lid-driven cavity at different Reynolds numbers. The planes $y = 0.5$, $x = 0.5$ and $z = 0.5$ are displayed from left to right; the Reynolds numbers concerned are Re = 100 (top), 400 (middle) and 1000 (bottom). The arrow indicates the associated direction of motion of the upper moving lid.

Figure 3.4 shows the results obtained with a grid containing $30 \times 30 \times 30$ points for Re = 100, 400 and 1000 which prove to be in very good agreement with those of Ding et al. [86], Jiang et al. [135], Ku et al. [153]. Figure 3.4d shows selected stream tubes for Re = 400 while Figure 3.5 visualises the corresponding tensor potential entries. Identification of the diagonal tensor entries as volume quantities and the off-diagonal entries as edge quantities (see Appendix A.2.4), when compared with the stress discretisation by Graves

[110], suggests a close relationship between the tensor entries and the stresses which opens up the opportunity to calculate the stresses from a_{ij}, an option that would justify the additional effort in calculating the tensor entries; this is left as a topic for future investigation. Finally Figure 3.6 displays the projected streamlines on the three mid-planar cross-sections for Re = 100, 400 and 1000; the results are consistent with, for example, those of Wang et al. [272].

(b) Unsteady flow and periodic boundary conditions

In relation to the DNS of viscous flow problems using a primitive variable formulation of the governing NS and continuity equations, the use of periodic geometries/domains - ones with boundaries that are periodic in each and every coordinate direction - can prove particularly advantageous. The pressure can be readily eliminated from the NS equations leading to a Poisson equation for the pressure which lends itself well to solution using pseudospectral methods since the pressure at the boundaries is easily specified. It results in the NS equations preserving the divergence free nature of the velocity field; as shown theoretically by Frisch [101] for what he equivalently terms a 3D periodicity cube.

Within the framework of the first integral the problem of having to solve a Poisson equation for the pressure can be avoided elegantly beginning with its traceless form, equations (3.68) and (3.70), from which the pressure field is completely absent.

Consider periodic boundary conditions for the streamfunction vector written as:

$$\Psi\left(x_1+L,x_2,x_3,t\right)=\Psi\left(x_1,x_2+L,x_3,t\right)=\Psi\left(x_1,x_2,x_3+L,t\right)=\Psi\left(x_1,x_2,x_3,t\right)\,,$$

together with multi-index notation for Greek letters, e.g. $\lambda=(\lambda_1,\lambda_2,\lambda_3)\in\mathbb{Z}^3$ with $\lambda^2=\lambda_p\lambda_p=\lambda_1^2+\lambda_2^2+\lambda_3^2$, followed by adopting a Fourier representation for both the stream function vector and the traceless tensor potential, namely:

$$\begin{aligned}\Psi_i(x_j,t)&=\sum_{\lambda}\Psi_i^{\lambda}(t)\exp\left(\mathrm{i}k_0\lambda_j x_j\right)\,,\\ \bar{a}_{pq}(x_j,t)&=\sum_{\lambda}\bar{a}_{pq}^{l}(t)\exp\left(\mathrm{i}k_0\lambda_j x_j\right)+\bar{a}_{pq}^{0}(t)x_m x_m\,,\end{aligned}$$

with $k_0=2\pi/L$, fulfilling periodic boundary conditions for the streamfunction vector automatically. The velocity field (3.37) then takes the form:

$$u_i=\sum_{\lambda}\underbrace{\mathrm{i}k_0\varepsilon_{ijn}\lambda_j\Psi_n^{\lambda}}_{u_i^{\lambda}}\exp\left(\mathrm{i}k_0\lambda_j x_j\right)\,.$$

Note that, the summation convention adopted for multi-indices is that summation is invoked only if the entire index, e.g. λ, occurs twice in a product; a single component of it,

e.g. λ_i, acts as a factor only and therefore does not affect summation.

Next, from equations (3.68), (3.70) the following set of equations result:

$$6\bar{a}^0_{ij} = -\varrho\left[u_i^{-\mu}u_j^{\mu} - \frac{u_p^{-\mu}u_p^{\mu}}{3}\delta_{ij}\right], \tag{3.116}$$

$$-k_0^2\lambda^2\bar{a}^{\lambda}_{ij} = -\varrho\left[u_i^{\lambda-\mu}u_j^{\mu} - \frac{u_p^{\lambda-\mu}u_p^{\mu}}{3}\delta_{ij}\right], \tag{3.117}$$

$$\varrho\dot{\Psi}^{\lambda}_n + \eta k_0^2\lambda^2\Psi^{\lambda}_n = k_0^2\varepsilon_{npq}\lambda_p\lambda_m\bar{a}^{\lambda}_{mq}\,. \tag{3.118}$$

Note that via equation (3.117) the coefficients of the traceless tensor potential can be expressed in terms of the coefficients of the velocity. Hence, all occurrences of $\bar{a}^{\lambda}_{mq}$ in equation (3.118) can be replaced, leading to:

$$\begin{aligned}\varrho\dot{\Psi}^{\lambda}_n + \eta k_0^2\lambda^2\Psi^{\lambda}_n &= \frac{\varrho}{\lambda^2}\varepsilon_{npq}\lambda_p\lambda_m\left[u_m^{\lambda-\mu}u_q^{\mu} - \frac{u_k^{\lambda-\mu}u_k^{\mu}}{3}\delta_{mq}\right]\\ &= \frac{\varrho}{\lambda^2}\left[\varepsilon_{npq}\lambda_p u_q^{\mu}\lambda_m u_m^{\lambda-\mu} - \underbrace{\varepsilon_{npq}\lambda_p\lambda_q}_{0}\frac{u_k^{\lambda-\mu}u_k^{\mu}}{3}\right].\end{aligned}$$

In addition, making use of the following identities:

$$\begin{aligned}\varepsilon_{npq}\lambda_p u_q^{\mu} &= \mathrm{i}k_0\varepsilon_{npq}\varepsilon_{qjk}\lambda_p\mu_j\Psi_k^{\mu} = \mathrm{i}k_0\left[\delta_{jn}\delta_{kp} - \delta_{jp}\delta_{kn}\right]\lambda_p\mu_j\Psi_k^{\mu} = \mathrm{i}k_0\left[\lambda_p\mu_n\Psi_p^{\mu} - \lambda_j\mu_j\Psi_n^{\mu}\right],\\ \lambda_i u_i^{\lambda-\mu} &= \mathrm{i}k_0\varepsilon_{ijq}\lambda_i(\lambda_j - \mu_j)\Psi_q^{\lambda-\mu} = -\mathrm{i}k_0\varepsilon_{ijq}\lambda_i\mu_j\Psi_q^{\lambda-\mu}\,,\end{aligned}$$

leads to:

$$\varrho\dot{\Psi}^{\lambda}_n + \eta k_0^2\lambda^2\Psi^{\lambda}_n - \frac{\varrho k_0^2}{\lambda^2}\left[\lambda_p\mu_n - \lambda_k\mu_k\delta_{np}\right]\varepsilon_{qij}\lambda_i\mu_j\Psi_q^{\lambda-\mu}\Psi_p^{\mu} = 0\,,$$

and therefore a set of quadratic equations for the coefficient functions $\Psi^{\lambda}_n(t)$. By truncating the set of equations after a finite number of modes corresponding to the values of the multi index λ, it can be solved numerically, to reveal the time evolution of the flow for a given initial state, e.g. a Taylor-Green vortex [101]; a topic of fundamental interest and for future exploration.

The attractiveness associated with periodic geometries has been mirrored in the interest shown in the use of pseudospectral methods for the solution of 3D viscous flows in non-periodic ones based on a primitive variable formulation – see, for example, the work of Ku et al. [153] – having at least one coordinate direction in which the boundaries are not periodic. The key related issues of deriving equations and boundary conditions for the pressure there which ensure satisfaction of the divergence free constraint on the velocity are comprehensively discussed by Tuckerman [260] with particular emphasis on the influence matrix method. However, as is rightly pointed out in the same article, the

solution of a Poisson equation for the pressure can be avoided completely by solving for the governing equations for the velocity and pressure fields together in a manner similar to the numerical scheme outlined in Appendix A.2.4, augmented with a suitably accurate temporal discretisation of the relevant terms in the governing equations. The same is clearly true if a solution based on a primitive variable formulation is preferred for unsteady flow in a periodic geometry.

3.5.4 Steady Stokes flow

The well known problem of the broadside translation of a thin disc through a viscous fluid [248], is considered. The unit disc $D = \{x \in \mathbb{R}^3 \,|\, (x_1^2 + x_2^2)^{1/2} \leq 1\,, x_3 = 0\}$ is located in the plane $P = \{x \in \mathbb{R}^3 \,|\, x_3 = 0\}$ and a constant disc velocity U_0 assumed, so that a steady unbounded and decaying velocity field (3.66) under conditions:

$$u_3(x) = U_0\,, x \in D \quad \text{and} \quad \partial_3 u_3(x) = 0\,, x \in P \setminus D\,,$$

is sought.

By assuming the potential energy density U to be zero and the vector $\mathcal{A}$ to be defined by:

$$\mathcal{A}_l = -\eta u_l - \partial_k \mathcal{D}(\tilde{a})_{kl} = \partial_k \tilde{a}_{kl} - \partial_k \mathcal{D}(\tilde{a})_{kl}\,, \tag{3.119}$$

with $\mathcal{D}(\tilde{a})$ denoting the diagonal part of the tensor $\tilde{a}$, then, from (3.60), it follows that $\partial_k \partial_k \mathcal{A}_l = 0$. The remaining gauge freedoms in (3.66) signify $\tilde{a}_{11} = \tilde{a}_{22} = \tilde{a}_{33}$, reducing the number of unknown fields from six to just four. As a consequence, using (3.62), the divergence of (3.119) results in $\partial_l \mathcal{A}_l = -\partial_l \partial_k \mathcal{D}(\tilde{a})_{kl} = -\partial_k \partial_k \tilde{a}_{11} = p$, providing a useful guide for the construction of a particular solution $\tilde{a}_{11}$; which is that the relationship:

$$\partial_k \partial_k x_l \mathcal{A}_l = x_l \partial_k \partial_k \mathcal{A}_l + 2\partial_l \mathcal{A}_l = 2p$$

facilitates the following decomposition involving an arbitrary harmonic field χ: $\tilde{a}_{11} = -\frac{1}{2} x_l \mathcal{A}_l + \chi$. What remains to be found is an overall solution procedure for obtaining four harmonic unknown fields $\mathcal{A}_l$ and χ for which the continuity equation is fulfilled identically and the velocity components are given by:

$$u_l = -\frac{1}{\eta}\left[\mathcal{A}_l + \partial_l\left(\chi - \frac{1}{2} x_k \mathcal{A}_k\right)\right]\,. \tag{3.120}$$

The pressure can be reconstructed subsequently via $p = \partial_l \mathcal{A}_l$. Equation (3.120) is equivalent to the Papkovich-Neuber representation known from elasticity theory [185] which generally allows for the analytical solution of various axis-symmetric problems as, for instance, shown by Rudge [209], Woodhouse and Goldstein [281] or Tran-Cong and Blake

[256]; moreover, it is closely related to the Clebsch transformation [154, 191, 225]. The above considerations lead to the representation of Papkovich and Neuber directly as a special case of the first integral of the NS equations, illustrating the elegance of this generalised theory.

Inspection of the flow geometry and imposition of the missing azimuthal dependency in the solution being sought leads to a reduced approach, that is equation (3.120) with $\mathcal{A}_1 = \mathcal{A}_2 = \chi = 0$; a manageable task utilising potential theory which can conveniently be written in cylindrical coordinates as:

$$\Delta\mathcal{A}(r,z) = \frac{\partial^2\mathcal{A}}{\partial r^2} + \frac{1}{r}\frac{\partial\mathcal{A}}{\partial r} + \frac{\partial^2\mathcal{A}}{\partial z^2} = 0 \quad \text{and} \quad \begin{cases} \mathcal{A} = -\eta U_0\,, & z = 0\,, \quad r \leq 1 \\ \frac{\partial\mathcal{A}}{\partial z} = 0\,, & z = 0\,, \quad r > 1 \end{cases} \tag{3.121}$$

involving $\mathcal{A} := \mathcal{A}_3$, $r := \sqrt{x_1^2 + x_2^2}$ and $z := x_3$. Problem (3.121) can be solved by either Hankel transform methods, see e.g. Tanzosh and Stone [248] and references therein, or through a Green's function representation combined with a clever reformulation of the fundamental singularity as provided by Ramm and Fabrikant [197]. A Hankel transform involving Bessel functions of the first kind leads to:

$$\mathcal{H}_\nu[\mathcal{A}] = \int_0^\infty \mathcal{A} r J_\nu\left(tr\right)\,\mathrm{d}r\,, \quad \mathcal{H}_\nu[\Delta\mathcal{A}] = \left(\frac{\partial^2}{\partial z^2} - t^2\right)\mathcal{H}_\nu[\mathcal{A}]\,, \tag{3.122}$$

resulting in $\mathcal{H}_\nu[\mathcal{A}] = A(t)\exp(-tz) + B(t)\exp(tz)$ with A and B independent of z. Symmetry considerations enable the calculation to be restricted to $z > 0$, giving $B = 0$ due to the decay condition. After performing an inverse Hankel transform:

$$\mathcal{A}(r,z) = \mathcal{H}_\nu^{-1}[A(t)\exp(-tz)] = \int_0^\infty A(t) t \exp\left(-tz\right) J_\nu\left(tr\right)\,\mathrm{d}t\,, \tag{3.123}$$

the boundary conditions on the right-hand side of (3.121) become:

$$\mathcal{A}(r,0) = \int_0^\infty A(t) t J_\nu\left(tr\right)\,\mathrm{d}t = -\eta U_0\,, \quad r \leq 1\,, \tag{3.124}$$

$$\frac{\partial\mathcal{A}}{\partial z}(r,0) = -\int_0^\infty A(t) t^2 J_\nu\left(tr\right)\,\mathrm{d}t = 0\,, \quad r > 1\,, \tag{3.125}$$

which, with reference to Gradshteyn and Ryzhik [109] (6.671 and 6.693) gives $A(t) = -2\eta U_0 \sin(t)/(\pi t^2)$ in the case of $\nu = 0$. Making use of integral calculus, see again Gradshteyn and Ryzhik [109] (6.752), an analytical solution of the form (3.123) is finally obtained:

$$\begin{aligned} \mathcal{A}(r,z) &= -\frac{2\eta U_0}{\pi}\int_0^\infty \frac{\sin(t)}{t}\exp\left(-tz\right) J_0\left(tr\right)\,\mathrm{d}t \\ &= -\frac{2\eta U_0}{\pi}\arcsin\left(\frac{2}{\sqrt{z^2 + (r+1)^2} + \sqrt{z^2 + (r-1)^2}}\right), \end{aligned} \tag{3.126}$$

enabling the velocity components to be written down via (3.120), leading to the same result as reported by Tanzosh and Stone [248].

3.5.5 Steady free surface flow in the presence of symmetries

So far, little attention has been paid to the steady, self-adjoint form of the first integral equations (3.73), complemented by the gauge (3.76)-(3.77), which can be derived via the variational principle (3.75) as described in Sec. 3.4.3. However, a description of the dynamics in terms of a scalar Lagrangian is convenient when a reformulation of the equations in terms of arbitrary curvilinear orthogonal coordinates q_1, q_2, q_3 is desired. The latter can facilitate a reduction in complexity for a respective flow problem if, for instance, the flow configuration is provided with a symmetry that implies independence of one of the three coordinates, say q_3, allowing for a formal reduction of the physical 3D problem to a 2D mathematical description. Such a curvilinear reformulation is provided below, essentially demonstrating that a reduction of the field equations to a very compact form, as manifest in (1.41)-(1.42) for Cartesian 2D flow, can similarly be retained in the presence of *arbitrary* symmetries; correspondingly, also a simple form of the dynamic boundary condition similar to (3.25) can be retained. The considerations below are of a general kind while a concrete example of coating flow over a sphere is presented later in Ch. 6.3.

Subsequently symmetry of the flow is assumed manifest by: (i) the independence of all fields from q_3 and (ii) the vanishing of all vector and tensor components in the q_3-direction. Both criteria are fulfilled by assuming the following form of the streamfunction vector and traceless tensor potential:

$$\Psi_k = \psi\left(q_\mu\right)\partial_k q_3\,,$$
$$\bar{a}_{ij} = a_{\mu\nu}\partial_i q_\mu \partial_j q_\nu\,,$$

where Greek indices run from 1 to 2 while Latin ones run from 1 to 3. Note that $a_{\nu\mu} = a_{\mu\nu}$ has to be fulfilled in order to ensure the symmetry of $\bar{a}_{ij}$. For ensuring that $\bar{a}_{ij}$ is traceless, the condition:

$$0 = \bar{a}_{ii} = a_{\mu\nu}\partial_i q_\mu \partial_i q_\nu = a_{\mu\nu}G_{\mu\nu}\,, \tag{3.127}$$

has to be fulfilled, where:

$$G_{\mu\nu} := \partial_i q_\mu \partial_i q_\nu\,, \tag{3.128}$$

is the metric tensor of the curvilinear coordinates. Note that due to the orthogonality of the coordinates, $G_{12} = G_{21} = 0$. Also it is convenient to define:

$$E_{\nu\mu} := \varepsilon_{ijk}\partial_i q_\nu \partial_j q_\mu \partial_k q_3\,. \tag{3.129}$$

Now, consider the general Lagrangian (3.74) with $f = 0$ and the simple substitution $\Phi := a_{nn}/6$, which is performed in order to draw parallels with the introductory description, Sec. 1.2.1(b), and the 2D case in Sec. 3.2, as will become clear below. With this substitution and the above assumptions and definitions, the Lagrangian (3.74) is re-written in Appendix A.2.3(a) in terms of the fields ψ, Φ, $a_{\mu\nu}$ depending on q_1, q_2 and their first order derivatives, as:

$$\ell = \varrho E_{\mu\lambda} E_{\nu\kappa} a_{\mu\nu} \partial_\lambda \psi \partial_\kappa \psi + \left[\eta E_{\nu\kappa} \partial_\kappa \psi - G_{\nu\kappa} \partial_\kappa \Phi\right] \left[2 \left\{G_{\mu\lambda} \partial_\lambda + \Delta q_\mu\right\} a_{\mu\nu} + \partial_\nu G_{\mu\lambda} a_{\mu\lambda}\right] . \tag{3.130}$$

Note that due to symmetry $a_{\nu\mu} = a_{\mu\nu}$ and condition (3.127) only two entries in the tensor potential are independent fields, say a_{11} and a_{12}, while $a_{21} = a_{12}$ and $a_{22} = -G_{11} a_{11}/G_{22}$. By variation with respect to the latter, two PDE result containing the streamfunction ψ and the potential Φ, but not containing a_{11} and a_{12}. These two Euler-Lagrange equations are therefore decoupled from the remaining two Euler-Lagrange equations resulting by variations with respect to ψ and Φ. As a consequence, the traceless tensor potential and related field equations do not have any effect on the physics and the number of equations and unknowns is effectively halved.

By means of (A.45) the first integral of the dynamic boundary condition can also be reduced as follows. By multiplying (3.82) with $\partial_i q_\mu$, the traceless tensor potential is eliminated, resulting in the two reduced boundary conditions ($\mu = 1, 2$):

$$\varepsilon_{ikl} \partial_i q_\mu \left[2\partial_k \Phi + \sigma n_k - U_k\right] \mathrm{d}x_l = 0 . \tag{3.131}$$

The vector $\mathrm{d}x_l$ in the above may be arbitrary, but since (3.131) is a triple product with two other vectors $\partial_i q_\mu$ and $2\partial_k \Phi + \sigma n_k - U_k$ having both no components in the q_3-direction, only the component of $\mathrm{d}x_l$ in the q_3-direction contributes to the condition. It can be assumed without loss of generality that: $\mathrm{d}x_l = \partial_l q_3 \mathrm{d}s$. By considering the decomposition $\partial_k \Phi = \partial_k q_\nu \partial_\nu \Phi$ of the Gradient of Φ into components in the q_ν-direction, and using the same decomposition for the normal vector and the auxiliary quantities U_k, namely:

$$n_k = \partial_k q_\nu \tilde{n}_\nu , \qquad U_k = \partial_k q_\nu \tilde{U}_\nu ,$$

respectively, equations (3.131) take the form:

$$\underbrace{\varepsilon_{ikl} \partial_i q_\mu \partial_k q_\nu \partial_l q_3}_{E_{\mu\nu}} \left[2\partial_\nu \Phi + \sigma \tilde{n}_\nu - \tilde{U}_\nu\right] \mathrm{d}s = 0 . \tag{3.132}$$

Since $E_{\mu\nu}$ is invertible according to $\det\left(E_{\mu\nu}\right) = E_{12}^2 \neq 0$, and $\mathrm{d}s$ arbitrary, one finally obtains:

$$2\partial_\nu \Phi + \sigma \tilde{n}_\nu - \tilde{U}_\nu = 0 , \tag{3.133}$$

as a reduced first integral of the dynamic boundary conditions. Regarding the computation of the auxiliary quantities U_i it can be assumed without loss of generality that the tangent vector $t_k^{(2)}$ is parallel to the gradient of q_3. Thus, derivatives with respect to s_2 vanish, allowing equation (A.38) to be fulfilled identically by:

$$U_j = \int U t_j^{(1)} \mathrm{d}s \,. \tag{3.134}$$

(a) Recovery of the 2D Cartesian formulation

The case of Cartesian 2D flow, $q_1 = x$, $q_2 = y$ and $q_3 = z$, implies $E_{\lambda\mu} = \varepsilon_{\lambda\mu}$ and $G_{\lambda\mu} = \delta_{\lambda\mu}$. According to (3.127), $0 = a_{11} + a_{22}$ has to be fullfilled, implying $a_{22} = -a_{11}$. The resulting Lagrangian (3.130) then reads as:

$$\begin{aligned}\ell = \varrho a_{11} \left[(\partial_y \psi)^2 - (\partial_x \psi)^2 \right] - 2\varrho a_{12} \partial_x \psi \partial_y \psi \\ + 2\left(\partial_x a_{11} + \partial_y a_{12}\right) \left[\eta \partial_y \psi - \partial_x \Phi\right] + 2\left(\partial_y a_{11} - \partial_x a_{12}\right) \left[\eta \partial_x \psi + \partial_y \Phi\right] \,.\end{aligned} \tag{3.135}$$

Variation of the action integral $I = \iint_A \ell \, \mathrm{d}x\mathrm{d}y$ with respect to a_{11} and a_{12} yields:

$$\delta a_{11}: \quad \varrho \left(u^2 - v^2\right) + 2\left\{\partial_x^2 - \partial_y^2\right\} \Phi - 4\eta \partial_x \partial_y \psi = 0 \,, \tag{3.136}$$

$$\delta a_{12}: \quad 2\varrho uv + 4\partial_x \partial_y \Phi + 2\eta \left\{\partial_x^2 - \partial_y^2\right\} \psi = 0 \,, \tag{3.137}$$

with the velocities u, v given by $u = \partial_y \psi$ and $v = -\partial_x \psi$, being equivalent to (1.41) in Ch. 1. Variations with respect to ψ and Φ deliver PDEs for a_{11} and a_{12} which are essentially decoupled from the above two equations and can therefore be neglected.

For a free surface given explicitly as $y = f(x)$ and a potential energy density $U = U(x, y)$ it follows, from (3.134), that $U_1 = \int U\left(x, f(x)\right) \mathrm{d}x$, $U_2 = \int U\left(x, f(x)\right) f'(x) \, \mathrm{d}x$, implying:

$$\frac{\partial \Phi}{\partial x} + \frac{\sigma f'(x)}{2\sqrt{1 + f'(x)^2}} - \frac{1}{2} \int U\left(x, f(x)\right) \mathrm{d}x = 0 \,, \tag{3.138}$$

$$\frac{\partial \Phi}{\partial y} - \frac{\sigma}{2\sqrt{1 + f'(x)^2}} - \frac{1}{2} \int U\left(x, f(x)\right) f'(x) \, \mathrm{d}x = 0 \,, \tag{3.139}$$

as the resulting form of the first integral of the dynamic boundary condition (3.133). Via (3.138), (3.139) the result (3.25) from Sec. 3.2.2(b) is reproduced.

(b) Axially symmetric flows in terms of spherical coordinates

Taking spherical coordinates $q_1 = r$, $q_2 = \vartheta$, $q_3 = \varphi$ and assuming that the flow does not depend on φ, implies $G_{11} = 1$, $G_{22} = r^{-2}$ and $E_{12} = -E_{21} = r^{-2}/\sin\vartheta$, while (3.127) yields: $0 = a_{\mu\nu} G_{\mu\nu} = a_{11} + a_{22}/r^2$, giving motivation for the substitutions:

$$a_{11} =: a_0 \,, \qquad a_{22} =: -r^2 a_0 \,, \qquad a_{12} =: r a_1 \,;$$

the latter substitution ensures that the two remaining potentials $a_{0,1}$ have the same physical dimension. Considering also that $\Delta q_1 = 2/r$ and $\Delta q_2 = \cot\vartheta/r^2$, the Lagrangian (3.130) takes the form:

$$\ell = \varrho a_0 \left[u_r^2 - u_\vartheta^2\right] + 2\varrho a_1 u_r u_\vartheta + 2\left[\eta u_\vartheta - \partial_r \Phi\right]\left[\frac{\partial_r(r^3 a_0)}{r^3} + \frac{\partial_\vartheta(a_1 \sin\vartheta)}{r\sin\vartheta}\right] + 2\left[\eta u_r - \frac{\partial_\vartheta \Phi}{r}\right]\left[\frac{\partial_r(r^3 a_1)}{r^3} - \frac{\partial_\vartheta(a_0 \sin\vartheta)}{r\sin\vartheta}\right], \tag{3.140}$$

with the radial and polar velocity given by:

$$u_r := \frac{\partial_\vartheta \psi}{r^2 \sin\vartheta}, \qquad u_\vartheta := -\frac{\partial_r \psi}{r\sin\vartheta}, \tag{3.141}$$

respectively. Variation of the action integral: $I = 2\pi \iint_A \ell\, r^2 \sin\vartheta \,\mathrm{d}r\mathrm{d}\vartheta$, with respect to $a_{0,1}$ delivers, after the necessary manipulation:

$$\delta a_0: \quad \frac{\varrho}{2}\left(u_r^2 - u_\vartheta^2\right) - r\partial_r\left[\frac{1}{r}\left(\eta u_r - \partial_r \Phi\right)\right] + \frac{1}{r}\partial_\vartheta\left[\eta u_\vartheta - \frac{\partial_\vartheta \Phi}{r}\right] = 0, \tag{3.142}$$

$$\delta a_1: \quad \varrho u_r u_\vartheta - r\partial_r\left[\frac{1}{r}\left(\eta u_\vartheta - \frac{\partial_\vartheta \Phi}{r}\right)\right] - \frac{1}{r}\partial_\vartheta\left[\eta u_r - \partial_r \Phi\right] = 0. \tag{3.143}$$

For a free surface given by $r = r_0 f(\vartheta)$, the first integral of the dynamic boundary condition (3.133) results, according to Appendix A.2.3(b), in:

$$\frac{\partial \Phi}{\partial r} + \frac{\sigma f(\vartheta)}{2\sqrt{f(\vartheta)^2 + f'(\vartheta)^2}} - \frac{\varrho g r_0^2}{4}\left[f(\vartheta)^2 \cos\vartheta + \int f(\vartheta)^2 \mathrm{d}\vartheta \sin\vartheta\right] = 0, \tag{3.144}$$

$$\frac{\partial \Phi}{\partial \vartheta} - \frac{\sigma r_0 f(\vartheta) f'(\vartheta)}{2\sqrt{f(\vartheta)^2 + f'(\vartheta)^2}} - \frac{\varrho g r_0^3}{4} f(\vartheta)\cos\vartheta \int f(\vartheta)^2 \mathrm{d}\vartheta = 0, \tag{3.145}$$

which, in its complexity, is comparable to the Cartesian 2D case (3.138), (3.139).

Part II

Construction of a New and Efficient Potential-based FE Solver – Implementation and Application

4 Discrete formulation and method of solution

In Part I of this work, potential field representations of the NS equations are presented, involving the establishment of a new and unconventional variational principle in Chapter 2 and facilitating their integration in Chapter 3. This Chapter explores ways in which the theoretical findings of Chapter 3 can be exploited in a computational sense for problems of engineering interest; with particular emphasis on the solution of 2D steady state flows.

In contrast to previous efforts at implementing such a solver with selected applications [230], the present work leads to a more *generic* approach based on a FE discretisation of the governing equations with the potential to solve large problems for arbitrary complex and non-smooth domains. Note that, the finite difference based numerical method utilised in Chapter 3 to solve the classical 3D lid-driven cavity problem was for proof of concept purposes, rather than an attempt to derive an optimal and flexible solution strategy for the solution of first integral formulations of fluid flow.

The new FE solver is established restricting attention to the solution of 2D flow problems but can be readily extended to 3D. It possesses a number of desirable attributes from both a numerical and computational perspective. These are:

(i) The method is formulated in a variational Galerkin type framework involving H^1 function spaces only and exhibits optimal rates of convergence.

(ii) Free surface boundary conditions and more generally stress boundary conditions are conveniently and efficiently handled by appropriate constraining of the potential field Φ; while the treatment of standard velocity boundary conditions is no more complicated than in the case of other classical FE formulations, namely velocity and pressure [55, 89, 137] or streamfunction and vorticity [103, 104, 107].

(iii) Implementation complexity is limited, in that many existing and established FE practices can be utilised; this compares well to the standard velocity-pressure formulation which in contrast requires the use of trial functions of different order.

(iv) The structure of the resulting linear systems is very favourable (being symmetric and positive definite); the conditioning is comparable to classical methods allowing solution by efficient and scalable linear solvers.

4.1 Preliminaries

Introduction of the streamfunction into the non-linear first integral equations (3.4a), (3.4b) and writing them in real-valued and non-dimensional form leads to:

$$\eta\left(\frac{\partial^2\Psi}{\partial y^2}-\frac{\partial^2\Psi}{\partial x^2}\right)-2\frac{\partial^2\Phi}{\partial x\partial y}+\mathrm{Re}\,\frac{\partial\Psi}{\partial x}\frac{\partial\Psi}{\partial y}=f_1\,, \tag{4.1a}$$

$$\frac{\partial^2\Phi}{\partial y^2}-\frac{\partial^2\Phi}{\partial x^2}+2\eta\frac{\partial^2\Psi}{\partial x\partial y}-\frac{\mathrm{Re}}{2}\left[\left(\frac{\partial\Psi}{\partial y}\right)^2-\left(\frac{\partial\Psi}{\partial x}\right)^2\right]=f_2\,, \tag{4.1b}$$

$$\Delta\Phi=p+\mathrm{Re}\frac{u^2}{2}+U\,, \tag{4.1c}$$

for some Reynolds number $\mathrm{Re}=\varrho U_0\ell/\eta$ with U_0 and ℓ being characteristic velocity and length scales. Ignoring equation (4.1c) for now, the remaining two equations, (4.1a) and (4.1b), can be expressed in a more convenient operator description given by $\mathcal{L}^2v+\mathcal{N}[v]=f$, for the solution vector $v=(\Psi,\Phi)$[1], where:

$$\mathcal{L}_{ij}:=\begin{bmatrix}\partial_y & -\partial_x\\ \partial_x & \partial_y\end{bmatrix}\,,\qquad \mathcal{N}_i[v]:=\frac{\mathrm{Re}}{2}\begin{bmatrix}2\partial_x\Psi\partial_y\Psi\\ (\partial_x\Psi)^2-(\partial_y\Psi)^2\end{bmatrix}\,, \tag{4.2}$$

At first sight a plausible weak Galerkin formulation of the above would have the form:

$$(\mathcal{L}^2v,w)_{0,\Omega}+(\mathcal{N}[v],w)_{0,\Omega}=(f,w)_{0,\Omega}\,,$$

for some test functions $w=(w_1,w_2)$. In order to ensure $H^1(\Omega)$ Hilbert space compliance, partial integration is required to eliminate second order derivatives; a subsequent solution would have to fulfil[2]:

Seek $v\in V$, so that for all $w\in W$ holds:

$$-(\mathcal{L}v,\mathcal{L}^*w)_{0,\Omega}+(\mathcal{L}v,\mathcal{L}^*(n)w)_{0,\partial\Omega}+(\mathcal{N}[v],w)_{0,\Omega}=(f,w)_{0,\Omega}\,, \tag{4.3}$$

where n is the outward pointing normal unit vector. Formally, in the above only first order derivatives of v remain; however, the corresponding "natural" boundary value problem requires the prescription $\mathcal{L}v(s)=g(s)$ for some boundary parametrisation s, which is not physically meaningful in the present context. In contrast, physically more meaningful velocity boundary conditions can be formed as a combination of Dirichlet and Neumann

[1] Note that, in the functional analytic context of Ch. 4 and of the corresponding Appendix B no use of explicit vector notation (involving vector arrows) is made. Instead the character and dimensionality of mathematical objects is given by means of their definition only, cf. [44].

[2] In formula (4.3), $\mathcal{L}^*$ denotes the adjoint of $\mathcal{L}$, while $\mathcal{L}^*(n)$ denotes the operator *symbol* of $\mathcal{L}^*$ with normal vector n as explained in Sec. B.2 of Appendix B; $(\cdot,\cdot)_0$ denotes the $L^2(\Omega)$ scalar product according to (B.3).

conditions purely for the streamfunction, namely:

$$\Psi(s) = g_1(s)\,, \tag{4.4}$$

$$\frac{\partial \Psi}{\partial \mathrm{n}} = g_2(s)\,. \tag{4.5}$$

The potential field Φ can be freely varied except for a finite number of constraints which guarantee a unique solution. Conditions (4.4), (4.5) have to be incorporated into the solution space V and thus represent *essential* boundary conditions for the system (4.3). The downside is that these conditions are of first order and over-constrain a classical $H^1(\Omega)$ FEM and thus necessitate at least a $H^2(\Omega)$ discretisation[3]; this level of complexity is undesirable.

The above is in contrast to the somewhat similar streamfunction-vorticity approach and its FE counterpart which was introduced by Ciarlet and Raviart [62] as a mixed method for solving the linear biharmonic equation and later extensively studied by several authors in the context of inertial flow problems [19, 103, 107], see also Bernardi et al. [30] and references therein. Using the streamfunction and the vorticity defined by the relations:

$$u = \frac{\partial \Psi}{\partial y}\,, \quad v = -\frac{\partial \Psi}{\partial x}\,, \quad \omega = \frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\,,$$

the steady 2D-NS equations can then be transformed into the system:

$$\begin{aligned} \Delta\omega - \mathrm{Re}\left(\frac{\partial \Psi}{\partial y}\frac{\partial \omega}{\partial x} - \frac{\partial \Psi}{\partial x}\frac{\partial \omega}{\partial y}\right) &= 0\,, \\ \Delta\Psi + \omega &= 0\,, \end{aligned}$$

with the same boundary conditions (4.4) (4.5). In terms of the solution spaces $\tilde{V}$, $\tilde{W}$ and the test spaces V, W:

$$\tilde{V} := \left\{ v \in H^1(\Omega) \,|\, v = g_1 \text{ on } \partial\Omega \right\}\,, \tag{4.6}$$

$$V := \left\{ v \in H^1(\Omega) \,|\, v = 0 \text{ on } \partial\Omega \right\}\,, \tag{4.7}$$

$$\tilde{W} = W := H^1(\Omega)\,, \tag{4.8}$$

a straightforward weak Galerkin formulation is given by the following:

Seek $\Psi \in \tilde{V}$, $\Phi \in \tilde{W}$, so that for all $v \in V$ and $w \in W$ the following holds:

$$\int_\Omega \nabla\omega \cdot \nabla v \,\mathrm{d}\Omega + \mathrm{Re} \int_\Omega v \left(\frac{\partial \Psi}{\partial y}\frac{\partial \omega}{\partial x} - \frac{\partial \Psi}{\partial x}\frac{\partial \omega}{\partial y}\right) \mathrm{d}\Omega - \int_{\partial\Omega} v \frac{\partial \omega}{\partial \mathrm{n}} \,\mathrm{d}\partial\Omega = 0\,, \tag{4.9}$$

$$\int_\Omega \nabla\Psi \cdot \nabla w \,\mathrm{d}\Omega - \int_\Omega w\omega \,\mathrm{d}\Omega - \int_{\partial\Omega} w g_2 \,\mathrm{d}\partial\Omega = 0\,, \tag{4.10}$$

[3] This disadvantage is not present in the case of a non-local, spectral (i.e. Fourier) discretisation as performed in [127, 230]

in which the boundary term in equation (4.9) can be omitted due to definition (4.7). Obviously the formulation (4.9), (4.10) allows for the convenient incorporation of natural Neumann conditions via the last term in (4.10), while Dirichlet boundary conditions are ensured by a restriction of the solution space (4.6), which is in contrast to the weak first integral approach (4.3). According to Ciarlet and Raviart [62] identical basis functions can be used for the discretised function spaces V, W as for Ψ, ω but the same may not be true for the function space for Ψ and Φ in the weak formulation (4.3)[4].

As shown later a simple and elegant remedy to the above problems is a least-squares stabilisation of the first integral equations, an idea which in the context of the Stokes part of system (4.1a), (4.1b) first appears in the work by Owens and Phillips [190]. By means of a least-squares approach the complication of mixed compatible element spaces for Φ and Ψ is avoided while a positive definite system is retained, even in the non-zero Reynolds number case. A variational least-squares formulation is only practical if the differential order is reduced appropriately, that is if the differential system is rewritten as a first-order PDE system; otherwise a discretisation in H^2-spaces becomes inevitable. A reformulation in terms of the velocity and the first-order derivatives of Φ is beneficial from another point of view: both velocity boundary conditions and free surface conditions can be incorporated as simple Dirichlet conditions, that is by constraining the appropriate function spaces, without disturbing the symmetry of the system. Such an approach is definitely promising with regard to attributes (i)-(iv) and forms the basis of the subsequent analysis.

4.2 First order system formulation

Since the overall goal is to construct a fairly general and flexible numerical method allowing for inertial effects and being potentially extendible to 3D, below the transformed real-valued version of the first integral equations (1.41)-(1.42) is discretised directly. Due to the close connection between these equations and a streamfunction / Airy stress function formulation as shown in Sec. 3.1, a review of the corresponding, though limited, literature [8, 38, 59, 69, 138, 252] was undertaken.

In the context of the present work a least squares finite element method, inspired by the contribution to the field of Bolton and Thatcher [38], Cassidy [59], Thatcher [252], proves adequate. This way the highly efficient semi-analytic Ritz method developed by Scholle et al. [230] for the integrated Stokes equations is complemented by a more flexible method, allowing for the incorporation of inertial effects and more general applicability. For a comprehensive review of least squares methods, including a special treatment of the Stokes and NS equations, the reader is referred to [37].

[4] Although this is not analysed in detail, numerical experiments suggest that identical function spaces for Ψ and Φ are not compatible and lead to oscillations.

The least-squares FEM (LSFEM) has gained great popularity for the numerical solution of flow problems, facilitating the use of simple equal order elements together with highly efficient multigrid solvers due to the symmetry and positive definiteness of the resulting system matrices [37]. Arguments of practicality suggest rewriting the tensor equations (1.41)-(1.42) in terms of velocity variables and first order derivatives of the Airy stress function leading to a system of four equations involving first order derivatives only, which is covered by the first order system least squares methodology.

The introduction of two field variables $\phi_1 = \partial_x \Phi$ and $\phi_2 = \partial_y \Phi$ together with the condition:

$$\frac{\partial \phi_1}{\partial y} - \frac{\partial \phi_2}{\partial x} = 0 \ , \tag{4.11}$$

enables equations (1.41), (1.42) to be written as:

$$-\varrho u_1 u_2 + \eta \left(\frac{\partial u_1}{\partial y} + \frac{\partial u_2}{\partial x} \right) - \frac{\partial \phi_1}{\partial y} - \frac{\partial \phi_2}{\partial x} = f_1 \ , \tag{4.12a}$$

$$\frac{\varrho}{2} \left(u_1^2 - u_2^2 \right) + \eta \left(\frac{\partial u_2}{\partial y} - \frac{\partial u_1}{\partial x} \right) + \frac{\partial \phi_1}{\partial x} - \frac{\partial \phi_2}{\partial y} = f_2 \ , \tag{4.12b}$$

$$\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} = 0 =: f_3 \ , \tag{4.12c}$$

$$\frac{\partial \phi_1}{\partial y} - \frac{\partial \phi_2}{\partial x} = 0 =: f_4 \ . \tag{4.12d}$$

Applying Newton linearisation to the above equations, with $\tilde{u}_1$ and $\tilde{u}_2$ denoting the velocity components of the respective previous iteration step, gives:

$$-\varrho(u_1 \tilde{u}_2 + \tilde{u}_1 u_2) + \eta \left(\frac{\partial u_1}{\partial y} + \frac{\partial u_2}{\partial x} \right) - \frac{\partial \phi_1}{\partial y} - \frac{\partial \phi_2}{\partial x} = f_1 - \varrho \tilde{u}_1 \tilde{u}_2 \ , \tag{4.13a}$$

$$\varrho(u_1 \tilde{u}_1 - u_2 \tilde{u}_2) + \eta \left(\frac{\partial u_2}{\partial y} - \frac{\partial u_1}{\partial x} \right) + \frac{\partial \phi_1}{\partial x} - \frac{\partial \phi_2}{\partial y} = f_2 + \frac{\varrho}{2} \left(\tilde{u}_1^2 - \tilde{u}_2^2 \right) \ , \tag{4.13b}$$

$$\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} = 0 \ , \tag{4.13c}$$

$$\frac{\partial \phi_1}{\partial y} - \frac{\partial \phi_2}{\partial x} = 0 \ , \tag{4.13d}$$

which can be non-dimensionalised in terms of l and U_0, with the potential fields ϕ_1, ϕ_2 scaled via ηU_0, and finally written in the following condensed form:

$$\mathcal{L}_{ij} u_j + 2\mathcal{N}_{ij}[\tilde{u}] u_j = f_i + \mathcal{N}_{ij}[\tilde{u}] \tilde{u}_j \, ; \tag{4.14}$$

the associated differential operators are defined as:

$$\mathcal{L}_{ij} := \begin{bmatrix} \partial_y & \partial_x & -\partial_y & -\partial_x \\ -\partial_x & \partial_y & \partial_x & -\partial_y \\ \partial_x & \partial_y & 0 & 0 \\ 0 & 0 & \partial_y & -\partial_x \end{bmatrix} , \qquad \mathcal{N}_{ij}[\tilde{u}] := \frac{\mathrm{Re}}{2} \begin{bmatrix} -\tilde{u}_2 & -\tilde{u}_1 & 0 & 0 \\ \tilde{u}_1 & -\tilde{u}_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} , \tag{4.15}$$

with $\mathrm{Re} = \varrho l U_0/\eta$ the Reynolds number and the solution vector given by:

$$u = (u_1, u_2, u_3, u_4) := (u_1, u_2, \phi_1, \phi_2)\,. \tag{4.16}$$

The operator $\mathcal{L}$ embodies the terms associated with the linear Stokes flow case, while $\mathcal{N}[\tilde{u}]$ contains terms related to inertial effects. In a similar way f describes possible external forces while $\mathcal{N}[\tilde{u}]\tilde{u}$ includes the terms which appear on the right-hand side from the linearisation process. Square brackets indicate the dependencies of the respective operators on other fields such as velocities from the previous iteration step. Obviously, in the chosen notation the exact non-linear system is given by:

$$\mathcal{L}_{ij} u_j + \mathcal{N}_{ij}[u] u_j = f_i\,, \tag{4.17}$$

while the Stokes flow case, $\mathrm{Re} \to 0$, becomes simply:

$$\mathcal{L}_{ij} u_j = f_i\,. \tag{4.18}$$

Note that, in subsequent analysis it is much more convenient to rename the solution vector of field variables as in (4.16) to allow for the compact notations (4.14) and (4.17).

Typically the system (4.17) is solved in a domain $\Omega \subset \mathbb{R}^2$ when either velocity boundary conditions, stress boundary conditions or a combination of both are prescribed on each part of the boundary $\partial\Omega = \bigcup_i \Gamma_i$; the boundary conditions take the form:

$$u_1 = g_{11}\,, \qquad u_2 = g_{12} \quad \text{on } \Gamma_1\,, \tag{4.19a}$$

$$\phi_1 = g_{21}\,, \qquad \phi_2 = g_{22} \quad \text{on } \Gamma_2\,, \tag{4.19b}$$

$$u_i\, n_i = g_{31}\,, \qquad \phi_i\, t_i = g_{32} \quad \text{on } \Gamma_3\,, \tag{4.19c}$$

$$u_i\, t_i = g_{41}\,, \qquad \phi_i\, n_i = g_{42} \quad \text{on } \Gamma_4\,, \tag{4.19d}$$

with position-depending boundary functions $g_{ij} = g_{ij}(s)$. Conditions (4.19a) correspond to fixed or moving walls, (4.19b) to the prescription of both, normal and tangential stresses and conditons (4.19c), (4.19d) to the prescription of either the normal velocity and the tangential stress components or vice versa; the latter being relevant for instance in the case of inflow and outflow conditions.

Seen individually, all of the above boundary conditions are physically meaningful; however, a mixture of them at different parts of the boundary frequently occurs in engineering applications as well, for instance in a channel problem with fixed walls and in/outflow conditions, cf. Sec. 5.3.1(c). Another application, particularly relevant in the context of the present work, is free surface flow where the position of the free boundary is an unknown quantity itself, cf. the considerations in Sec. 5.1.2: Here, typically an iteration process is employed in which a sequence of problems with fixed boundaries is solved with (4.19a) prescribed at the walls and (4.19b), determined by (3.25), at the free surface.

A priori estimates for the analytic solution of the problem are necessary as a basis for a convergence analysis, that is estimates of the solution against the right-hand side of the field equations as well as the boundary conditions are sought. For a rather general class of elliptic boundary value problems, the present one included, such estimates can be obtained from the famous theory of Agmon, Douglis and Nirenberg [2], [3], abbreviated as ADN-theory (Sec. 4.3.1). The specific norms involved in the ADN-estimate restrict the reasonable choice of norms for the least-squares minimisation: in the present case the first integral system is classified as an elliptic system of Petrovskii type which proves to be particularly useful as it allows for an efficient minimisation in the $L^2(\Omega)$-norm.

A convergence analysis of the full non-linear problem is currently not feasible; however, a full analysis of the tractable linear (Stokes) problem is undertaken, the details of which are provided below. For convenience, boundary conditions are limited to the case where *one* of the conditions (4.19a)-(4.19d) is prescribed at the *whole* boundary; for this simplified case the complementing condition due to Agmon et al. [2, 3] can be verified and a priori estimates derived. Also complications in the error analysis associated with where different boundary conditions meet are avoided; these are investigated in Sec. 4.3.4. At this point it needs to be clarified whether a unique solution of the system (4.17) with the boundary conditions mentioned can be obtained or whether additional constraints for the potential fields ϕ_1 and ϕ_2 may be required.

4.3 Weak form and convergence analysis of the Stokes problem

4.3.1 A priori estimates for the linear equations

In order to derive a priori estimates for system (4.18) with boundary conditions (4.19) that are relevant for the analysis of a corresponding least-squares method, the elliptic regularity ADN-theory is employed, following in the footsteps of Bochev and Gunzburger [37]. A review of ADN-theory together with fundamental definitions is provided in detail in Appendix B.2.

A major outcome of the above review represents Theorem B.33 which can assist in deriving a priori error bounds for the first integral formulation of Sec. 4.2, when as a first step the corresponding preconditions have been verified. Consider the field equations (4.18) with boundary conditions (4.19) in dimensionless form; this is obviously a boundary value problem of the form (B.20), (B.21) with a differential field operator symbol given by:

$$\mathcal{L}(x,\vec{\xi}) = \mathcal{L}(\vec{\xi}) := \begin{bmatrix} \xi_2 & \xi_1 & \xi_2 & \xi_1 \\ -\xi_1 & \xi_2 & -\xi_1 & \xi_2 \\ \xi_1 & \xi_2 & 0 & 0 \\ 0 & 0 & \xi_2 & -\xi_1 \end{bmatrix} . \tag{4.20}$$

Now, the two sets of integer weights for the equations and unknowns are defined as:

$$(s_1,\, s_2,\, s_3,\, s_4) = (0,\, 0,\, 0,\, 0)\,, \tag{4.21}$$

$$(t_1,\, t_2,\, t_3,\, t_4) = (1,\, 1,\, 1,\, 1)\,, \tag{4.22}$$

which directly allow for the verification of:

$$\deg \mathcal{L}_{ij}(\vec{\xi}) \leq 1 = s_i + t_j\,, \quad \text{for all } i,j = 1,\ldots,4\,.$$

It is easy to see that the principal part of $\mathcal{L}$ equals the original operator, i.e. $\mathcal{L}^p(\vec{\xi}) = \mathcal{L}(\vec{\xi})$, and that:

$$\det \mathcal{L}^p(\vec{\xi}) = (\xi_1^2 + \xi_2^2)^2 = |\vec{\xi}|^4 \neq 0\,, \quad \text{for all } \vec{\xi} \neq 0\,, \tag{4.23}$$

which, together, establishes ADN-ellipticity of the operator $\mathcal{L}$ according to Def. B.29; by the choice of integer weights the system is even elliptic in the sense of Petrovskii, see Def. B.31. Moreover, the identity (4.23) gives $\deg(\det \mathcal{L}^p(x,\vec{\xi})) = 4 =: 2m$ and for $m = 2$ and $C = 1$ condition (B.24) is also verified:

$$|\vec{\xi}|^4 \leq |\det \mathcal{L}^p(x,\vec{\xi})| = |\vec{\xi}|^4 \leq |\vec{\xi}|^4\,,$$

and uniform ellipticity established. The fulfilment of the Supplementary Condition B.30 and thus the regularity of the system is seen by recognising that the polynomial:

$$\begin{aligned}\det \mathcal{L}^p(\vec{\xi} + \tau\vec{\xi}') = |\vec{\xi} + \tau\vec{\xi}'|^4 &= \left(|\vec{\xi}|^2 + \tau^2|\vec{\xi}'|^2\right)^2 \\ &= \left(|\vec{\xi}| + \mathrm{i}\tau|\vec{\xi}'|\right)^2 \left(|\vec{\xi}| - \mathrm{i}\tau|\vec{\xi}'|\right)^2\,,\end{aligned}$$

has exactly $m = 2$ identical roots with positive imaginary part in the complex variable τ.

Next consider the four relevant differential boundary operators, see Appendix B.2, associated with (4.19), namely:

$$\mathcal{B}_1(x,\vec{\xi}) := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}\,, \quad \mathcal{B}_2(x,\vec{\xi}) := \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\,, \tag{4.24}$$

$$\mathcal{B}_3(x,\vec{\xi}) := \begin{bmatrix} n_1(x) & n_2(x) & 0 & 0 \\ 0 & 0 & t_1(x) & t_2(x) \end{bmatrix}\,, \tag{4.25}$$

$$\mathcal{B}_4(x,\vec{\xi}) := \begin{bmatrix} t_1(x) & t_2(x) & 0 & 0 \\ 0 & 0 & n_1(x) & n_2(x) \end{bmatrix}\,, \tag{4.26}$$

noting that for the boundary equation weights $(r_1,\, r_2) = (-1,\, -1)$ and all variants of the boundary operators, the following equality holds:

$$\deg \mathcal{B}_{lj}(x,\vec{\xi}) = 0 = r_l + t_j\,, \quad \text{for all } l = 1,2,\ j = 1,\ldots,4\,.$$

It remains to show "compatibility" between boundary conditions and differential field operator, i.e. that the Complementing Condition B.32, holds. For this purpose consider the normal unit vector $\vec{n}$ to the boundary, an arbitrary tangential counterpart $\vec{\xi} \perp \vec{n}$, $\|\vec{\xi}\| = 1$ and the polynomial:

$$M^+(x, \vec{\xi}, \tau) = \prod_{k=1}^{m} \left(\tau - \tau_k^+(x, \vec{\xi}) \right) = (\tau - \mathrm{i})^2 \,.$$

Moreover, the cofactor matrix $\mathcal{L}'$ is identified as:

$$\mathcal{L}'(\vec{\xi}) := \det(\mathcal{L}^p)\mathcal{L}^{-p} = \begin{bmatrix} \xi_2(\xi_2^2 - \xi_1^2) & -2\xi_1\xi_2^2 & \xi_1(\xi_1^2 + \xi_2^2) & -\xi_1(\xi_1^2 + \xi_2^2) \\ \xi_1(\xi_1^2 - \xi_2^2) & 2\xi_1^2\xi_2 & \xi_2(\xi_1^2 + \xi_2^2) & \xi_1(\xi_1^2 + \xi_2^2) \\ 2\xi_1^2\xi_2 & \xi_1(\xi_2^2 - \xi_1^2) & -\xi_1(\xi_1^2 + \xi_2^2) & \xi_2(\xi_1^2 + \xi_2^2) \\ 2\xi_1\xi_2^2 & \xi_2(\xi_2^2 - \xi_1^2) & -\xi_2(\xi_1^2 + \xi_2^2) & -\xi_1(\xi_1^2 + \xi_2^2) \end{bmatrix},$$

and it is observed that the principal parts $\mathcal{B}^p$ of the boundary operators in (4.24)-(4.26) trivially equal the corresponding primary operators $\mathcal{B}$. With the definition $\vec{\eta} := \vec{\xi} + \tau\vec{n}$, according to (B.25) an auxiliary matrix can be constructed for each of the three boundary operators in (4.24), (4.25), namely:

$$\sum_{j=1}^{4} \mathcal{B}_{lj}^{p,1}(x, \vec{\eta})\mathcal{L}'_{jk}(x, \vec{\eta}) = \mathcal{L}'_{l,k}(x, \vec{\eta}) \,, \quad l = 1, 2; \; k = 1, \ldots, 4 \tag{4.27}$$

$$\sum_{j=1}^{4} \mathcal{B}_{lj}^{p,2}(x, \vec{\eta})\mathcal{L}'_{jk}(x, \vec{\eta}) = \mathcal{L}'_{l+2,k}(x, \vec{\eta}) \,, \quad l = 1, 2; \; k = 1, \ldots, 4 \tag{4.28}$$

$$\sum_{j=1}^{4} \mathcal{B}_{lj}^{p,3}(x, \vec{\eta})\mathcal{L}'_{jk}(x, \vec{\eta}) = \begin{cases} n_q\mathcal{L}_{qk}(x, \vec{\eta}) \,, & l = 1; \; k = 1, \ldots, 4 \\ t_q\mathcal{L}_{qk}(x, \vec{\eta}) \,, & l = 2; \; k = 1, \ldots, 4 \,. \end{cases} \tag{4.29}$$

Any non-trivial linear combination of rows in the matrices (4.27)-(4.29) has elements which are cubic in τ and these cannot be integer multiples of the τ-quadratic polynomial M^+. Also these rows are linearly independent and therefore the complementing condition is satisfied.

Having shown the ellipticity of (4.18) in combination with boundary conditions of the type (4.19a)-(4.19d), Theorem B.34 guarantees the Fredholm property for the operator $(\mathcal{L}, \mathcal{B})$ and thus a finite-dimensional null-space. Thus, at most finitely many constraints have to be formulated in order to achieve a uniquely solvable system. If these constraints are found, the solution of the constrained system exists if and only if a solution of the corresponding Stokes problem in primitive variables exists which can be assumed without loss of generality in the present context; then the resulting velocities coincide. For further details concerning the existence and uniqueness theory of the Stokes equations the reader is referred to Hackbusch [119]. However, the desired constraints depend on the precise boundary conditions and can be found by simple calculations, cf. [251]:

Lemma 4.1. (Unique solution for Stokes flow)

(i) Equations (4.18) *with boundary conditions* (4.19a) *have a unique solution u if additionally the following constraints are fulfilled:*

$$\Lambda_1 u := \int_\Omega \phi_1 \,\mathrm{d}\Omega = 0\,, \quad \Lambda_2 u := \int_\Omega \phi_2 \,\mathrm{d}\Omega = 0\,, \quad \Lambda_3 u := \int_\Omega \frac{\partial \phi_1}{\partial x} \,\mathrm{d}\Omega = 0\,. \tag{4.30}$$

(ii) Equations (4.18) *with boundary conditions* (4.19b) *have a unique solution u if additionally the following constraints are fulfilled:*

$$\Lambda_1 u := \int_\Omega u_1 \,\mathrm{d}\Omega = 0\,, \quad \Lambda_2 u := \int_\Omega u_2 \,\mathrm{d}\Omega = 0\,, \quad \Lambda_3 u := \int_\Omega \frac{\partial u_1}{\partial x} \,\mathrm{d}\Omega = 0\,. \tag{4.31}$$

(iii) Equations (4.18) *with boundary conditions either of the form* (4.19c) *or* (4.19d) *have a unique solution u without additional constraints.*

Proof. To show uniqueness of the solution of (4.18), it is sufficient to consider the homogeneous equations with homogeneous boundary conditions (4.19). Equations (4.18) in terms of the streamfunction Ψ and the potential field Φ, the existence of which is ensured by (4.12c) and (4.12d), can be written as:

$$\frac{\partial^2 \Psi}{\partial y^2} - \frac{\partial^2 \Psi}{\partial x^2} + 2\frac{\partial^2 \Phi}{\partial x \partial y} = 0\,, \tag{4.32}$$

$$\frac{\partial^2 \Phi}{\partial y^2} - \frac{\partial^2 \Phi}{\partial x^2} - 2\frac{\partial^2 \Psi}{\partial x \partial y} = 0\,. \tag{4.33}$$

For sufficiently smooth field variables Ψ and Φ, the combinations $\{\partial_{yy} - \partial_{xx}\}(4.32) - 2\partial_{xy}(4.33)$ and $\{\partial_{yy} - \partial_{xx}\}(4.33) + 2\partial_{xy}(4.32)$ lead to two biharmonic equations:

$$\Delta^2 \Psi = 0\,, \qquad \Delta^2 \Phi = 0\,.$$

With regard to (i), clearly the biharmonicity of Ψ in combination with prescribed velocities (4.19a) specifies Ψ up to a constant and therefore the velocities uniquely in the whole domain [106]; for homogeneous boundary conditions then $u_1 = u_2 = 0$ in Ω. Now, from equations (4.12a) and (4.12d):

$$\frac{\partial \phi_1}{\partial y} + \frac{\partial \phi_2}{\partial x} = 0\,, \qquad \frac{\partial \phi_1}{\partial y} - \frac{\partial \phi_2}{\partial x} = 0\,,$$

thus giving $\partial_y \phi_1 = \partial_x \phi_2 = 0$ and the reduced dependencies $\phi_1 = \phi_1(x)$, $\phi_2 = \phi_2(y)$, respectively. Furthermore (4.12b) gives $\partial_y \phi_2 - \partial_x \phi_1 = 0$ and, considering the mentioned dependencies, the constancy of these derivatives is established, i.e., $\partial_y \phi_2 = \partial_x \phi_1 = c_1$. The three constants in the solution representation:

$$\phi_1(x) = c_1 x + c_2\,, \qquad \phi_2(y) = c_1 y + c_3\,,$$

are uniquely determined by the constraints (4.30) and it is obvious that the prescription of a zero mean derivative with respect to x can be replaced by constraining an arbitrary linear combination of both partial derivatives. With regard to (ii) the proof is completely analogous and therefore omitted.

Constraint (iii) is technically more involved and can in principle be shown by a transformation to boundary-fitted orthogonal coordinates. The most general case of mixed boundary conditions, involving all of the four mentioned variants on different parts of the boundary, is not considered.

□

In the following, constraints of the form (4.30), (4.31) are simply denoted by $\Lambda u = c$ with a linear operator $\Lambda : H^1(\Omega)^4 \to \mathbb{R}^\alpha$ where, for instance, $\alpha = 3$ in the case of (4.30), (4.31) and $\alpha = 0$ otherwise. Note that Λ is $H^1(\Omega)$-continuous in the sense that:

$$\sum_{j=1}^{\alpha} \Lambda_j u \Lambda_j v \leq c \sum_{j=1}^{4} \|u_j\|_1 \|v_j\|_1 \,, \quad \text{for all } u, v \in H^1(\Omega)^4 \,, \tag{4.34}$$

which becomes important later. For subsequent analysis the complete and uniquely solvable system (4.18) together with one of the admissible boundary conditions (4.19a)-(4.19c) and associated constraints from Lemma 4.1 is denoted as:

$$\begin{cases} \mathcal{L}u &= f \,, \\ \mathcal{B}u &= g \,, \\ \Lambda u &= c \,. \end{cases} \tag{4.35}$$

Finally from Theorem B.28, a Schauder estimate can be derived by specifying $t' = \max t_j = 1$ and $q \geq r' = \max(0, \max r_l + 1) = 0$. Additionally, considering Corollary B.35, yields the following a priori estimate [37]:

Theorem 4.2. *Consider the elliptic boundary value problem* (B.20), (B.21) *with the differential field operator symbol* (4.20) *and one of the three boundary operator symbols* (4.24)-(4.26). *Let* $q \geq 0$ *and assume that* Ω *is a bounded domain of class* C^{q+1}. *If* $f_i \in H^q(\Omega)$, $g_l \in H^{q+1/2}(\partial\Omega)$, *then:*

(i) every solution $u \in L^2(\Omega)^4$ *in fact belongs to* $H^{q+1}(\Omega)^4$;

(ii) there is a positive constant C, *independent of* u, f *and* g, *such that, for every solution* $u \in \prod_{j=1}^4 H^{q+1}(\Omega)$,

$$\sum_{j=1}^{4} \|u_j\|_{q+1} \leq C \left(\sum_{i=1}^{4} \|f_i\|_q + \sum_{l=1}^{2} \|g_l\|_{q+1/2,\partial\Omega} + \sum_{j=1}^{3} |\Lambda_j u| \right) . \tag{4.36}$$

Part (i) in the above theorem is called a "regularity shift" in the literature [118]; that is, a higher regularity in the right-hand sides f and g of the elliptic problem leads to an increased regularity of the solution u. Note that, in the classical theory of elliptic boundary value problems very strong assumptions are made on the smoothness of the domain boundary, here $\partial\Omega \in C^{q+1}$, in order to guarantee the regularity shift and thus a priori estimates of the form (4.36). For some problems of practical relevance involving polygonal boundary parts with angular corners or cusps, such an assumption is inappropriate and can be weakened, although not for general ADN-elliptic systems. A glance at the vast literature on Stokes equations in primitive variables, see [44, 107, 118] and references therein, reveals that in the case of Dirichlet boundary conditions a velocity solution $u \in H^1(\Omega)$ is obtained under the considerably weaker assumption of a Hölder-continuous[5] boundary $\partial\Omega \in C^{0,1}$ (Proposition 12.2.14 in [118]) which is expected to translate to the above case when $q = 0$.

However, as will be seen later, relevant qualitative error estimates require a solution regularity of $H^2(\Omega)$ which is generally not available in non-smooth domains [115]: for instance the nature of a conical boundary point depends on the interior angle [32]. For convex polygonal domains the Stokes equations in primitive variables admit a $H^2(\Omega)$-solution, [118] Proposition 12.2.19, while interior angles greater than π introduce more severe complications: here the solution may generally involve singularities not being captured by $H^2(\Omega)$.

In order to analyse non-smooth domains in detail, it is not reasonable any more to work with a homogeneous error measure, instead weighted Sobolev norms are introduced with weighting factors dependant on the distance to conical points. The corresponding weighted Sobolev spaces were first introduced by Kondrat'ev [149] and since then the theory of elliptic systems in non-smooth domains has been further developed by [73, 115, 206]. The subsequent analysis is mostly confined to sufficiently smooth domains while such irregular domains are treated in Sec. 4.3.4.

Remark 4.3. Throughout the above analysis, for instance in inequality (4.36) and in the definition of the function spaces (B.27), the sum and product symbols and thereby the dimensionality of objects like u, f and g were explicitly written down to avoid any confusion. In the literature, see [215, 275], a shorter notation is often preferred due to better legibility and the fact that the dimensionality of the objects (norms, spaces, etc.) used is clear in the context of and with their definition. In the subsequent analysis, the sum symbols in expressions like (4.36) are omitted for convenience while the respective norms are always consistently interpreted as depending on the object inside the norm brackets. The dimensionality exponent n in expressions like $H^1(\Omega)^n$, referring to a n-dimensional product space, is omitted in favour of better legibility, unless it is explicitly important in

[5] For the definition of Hölder continuity, $C^{k,\gamma}$, see Def. 3.2.8 in [118].

the context; thus, if $u \in H^1(\Omega)$ is written for a *vectorial* quantity u, the convention is that all components u_j are in $H^1(\Omega)$.

4.3.2 Weak variational formulation

(a) Definition of the weak variational problem

As shown in the previous section, representing the first integral equations as a first order system allows the identification of norm indices t_j, s_i and r_l such that the requirements of ADN-theorem B.33 are fulfilled; especially the ellipticity of the differential operator $\mathcal{L}_{ij}$, see (4.18), and the compatibility of the boundary operators $\mathcal{B}_{lj}$, see (4.19), in the form of the complementing condition Definition B.32 are ensured by this choice of indices. As a consequence, ADN-theory provides the practical a priori estimate (4.36) clearly indicating with respect to which norm a defect minimisation is reasonable.

Following the convention of Remark 4.3, the a priori estimate (4.36) suggests a minimisation of the functional:

$$\varrho_q^2(v-u) = \|\mathcal{L}v - f\|_q^2 + \|\mathcal{B}v - g\|_{q+1/2,\partial\Omega}^2 + |\Lambda v - c|\,, \quad q \geq 0\,, \tag{4.37}$$

$$\varrho_q^2(u_h - u) = \min_{v \in V} \varrho_q^2(v-u)\,, \tag{4.38}$$

in suitable function spaces, $v \in V$, which forces the error norm $\|u - u_h\|_{q+1} \leq \sqrt{2} C \varrho_q(u_h)$ to be small and is at least sufficient to obtain a convergent method [275]. In view of the term $\mathcal{L}v - f$, the simplest numerical method results for $q = 0$ leading in principle to a convenient $L^2(\Omega)$-minimisation of the residual equations and to an error $\|u - u_h\|_1$ which is bounded from above. However, the additional term $\|\mathcal{B}u_h - g\|_{1/2,\partial\Omega}$ cannot be calculated numerically without considerable difficulty[6] as noted by Wendland [275], Ch. 8.2. A simple remedy, replacing the above term by $\|\mathcal{B}u_h - g\|_{1,\partial\Omega} \leq \|\mathcal{B}u_h - g\|_{1/2,\partial\Omega}$, leads to the following convergence result, Wendland [275], for general homogeneous elliptic systems in the plane:

Lemma 4.4. *For a mesh-dependent family of approximation spaces V_h let $\bigcup_{h \leq h_0} V_h$ be dense in $H^1(\Omega)$ and let the corresponding functions on $\partial\Omega$ be dense in $H^1(\partial\Omega)$ for every $h_0 > 0$. Furthermore, let Λ_j be continuous on $H^1(\Omega)$, then:*

$$\lim_{h \to 0} \|u - u_h\|_1 = 0\,.$$

However, the above result does not make a statement as to the *order* of convergence which is a decisive criterion by which to assess a method. Instead, a further analysis [275] using nodal Lagrange basis functions with piece-wise k-th order polynomials from

[6] Compare with the formal definition of fractional Sobolev norms in the Appendix, Definition B.12.

the space $V \subset H^m(\Omega) \cap C^{m-1}(\overline{\Omega})$, gives, assuming that the exact solution u is at least in $H^{m+1}(\Omega)$, the following refined and more general result:

Lemma 4.5. *The least-squares method* (4.37), (4.38) *with* $q = 0$ *and* $\|\cdot\|_{1/2,\partial\Omega}$ *replaced by* $\|\cdot\|_{1,\partial\Omega}$ *converges for* $m \geq 2$ *and* $3/2 \leq k \leq m+1$ *as:*

$$\|u - u_h\|_1 \leq ch^{k-3/2}\|u\|_k \,.$$

The above inequality has two obvious disadvantages: the simplest case of a continuous and piece-wise linear FE analysis is not admissible with this method; moreover, the order of convergence is suboptimal since it is of order 1/2 less than that of the corresponding FE approximation. Both disadvantages disappear when an even simpler method, but involving an h-dependent weighting of the norms, is used. By applying a least squares method with a penalty in the boundary condition, which leads to the minimisation of the defect functional:

$$\begin{aligned} \varrho_0^2(v-u) &= \|\mathcal{L}v - f\|_0^2 + h^{-1}\|\mathcal{B}v - g\|_{0,\partial\Omega}^2 + |\Lambda v - c| \,, \\ \varrho_0^2(u_h - u) &= \min_{v \in V} \varrho_0^2(v-u) \,, \end{aligned}$$

in the chosen unrestricted trial subspaces, Wendland [275] achieved the optimal convergence:

$$\|u - u_h\|_0 + h\|u - u_h\|_1 \leq ch^2\|u\|_2 \,, \tag{4.39}$$

if linear elements are employed. However, convergence was proved only in the case where first-order Petrovskii systems of the form (4.35) are solvable for all right sides f, g. In the general case the same order of convergence was achieved by adjoining a finite number of suitable unknowns into equations (4.35) and approximating these in the corresponding modified method. Such an approach requires that the solvability conditions for (4.35) are known approximately.

In the present work another approach is utilised, along the lines of Saranent and Wendland [215], where the boundary conditions are approximated point-wise at the boundary nodes of the domain decomposition. This procedure leads to an error estimate consistent with (4.39) but is more practical in dealing with non-differential boundary conditions, especially Dirichlet conditions of the form (4.24), (4.25); although *restricted* trial spaces are introduced into the analysis.

What follows is mainly based on the work of Saranent and Wendland [215], [275] where a general theory for the least-squares FE approximation of first-order elliptic systems has been developed in the context of linear elements. The results are applied to the first integral equations (4.35) and generalised to higher order FEs with curved element edges,

a possibility commonly used in the context of isoparametric FEs. In contrast to [275] the present approach allows proof of convergence without any auxiliary variables which is mainly due to the use of a corresponding Galerkin approximation result for the second-order elliptic operator $\mathcal{L}^*\mathcal{L}$ with appropriate boundary conditions.

For the procedure with essential Dirichlet conditions, firstly the approximation in the case of homogeneous boundary conditions is considered:

$$\left\{ \begin{array}{rcl} \mathcal{L}u & = & f\,, \\ \mathcal{B}u & = & 0\,, \\ \Lambda u & = & 0\,; \end{array} \right. \tag{4.40}$$

non-homogeneous conditions are left to Sec. 4.3.3(b). In this case, the least-squares principle can be formulated with restricted but identical solution and trial spaces, that is:

Find $u_h \in V_h$, such that:

$$\mathcal{J}(u_h) := \varrho^2(u_h - u) = \|\mathcal{L}u_h - f\|_0^2 + |\Lambda u_h - c|^2 \leq \mathcal{J}(v) \quad \forall\, v \in V_h\,. \tag{4.41}$$

The restricted function space V_h is obviously related to the mesh and the particular FE discretisation; further notation has to be introduced at this point which is kept to a minimum. Let $\mathcal{T}_h$, $0 < h \leq h_0$ define a family of admissible domain decompositions of $\overline{\Omega}$, the requirements of which will be clarified in Sec. 5.1.3(b). This decomposition into arbitrary straight-bounded elements may allow for curved edges at the domain boundary but will, in general, not be enough to capture $\overline{\Omega}$ exactly; however, the closed domain covered by the decomposition, subsequently denoted by $\overline{\Omega}_h = \bigcup_{T\in\mathcal{T}_h}$, in the limit converges to $\overline{\Omega} = \lim_{h\to 0} \overline{\Omega}_h$. In principle, approximations of the field variables u_i are defined on the predefined domain decomposition but without loss of generality this approximation may be continuously extended to the boundary by the zero solution while overlapping parts may be cut off, so that any $u_i \in H^1(\Omega_h)$ uniquely represents a modified $u_i \in H^1(\Omega)$. In respect thereof consider finite dimensional subspaces of $H^1(\Omega)^4$, given by:

$$V_h := \left\{ v \in H^1(\Omega)^4 \cap C^0(\overline{\Omega})^4 \,\middle|\, \mathcal{B}(x)v(x) = 0 \text{ for all } x \in \partial\Omega_h \right\}\,, \tag{4.42}$$

such that the elements of V_h are continuous in $\overline{\Omega}$ as well as piece-wise smooth with respect to the decomposition $\mathcal{T}_h$; additionally they satisfy the boundary conditions:

$$\mathcal{B}(x)v(x) = 0\,, \tag{4.43}$$

for all boundary nodes $x \in \partial\Omega_h$ of the decomposition with one of the boundary operators given in (4.24)-(4.26). The point-wise condition (4.43) implies a useful estimate of the boundary integral of $\mathcal{B}v$, $v \in V_h$ as demonstrated in Sec. 5.1.3(d) for the FE spaces under consideration, which is treated as an assumption for the time being:

Assumption 4.6. *Assume an isoparametric FE approximation involving polynomials of order k, then there exists $h_1 > 0$ and $c > 0$ such that for all $v \in V_h$, $0 < h \leq h_1$ the following holds:*

$$\|\mathcal{B}v\|_{0,\partial\Omega} \leq ch^{(2k+1)/2}\|v\|_{1,\Omega}\,, \tag{4.44}$$

$$\|\mathcal{B}v\|_{1,\partial\Omega} \leq ch^{(2k-1)/2}\|v\|_{1,\Omega}\,. \tag{4.45}$$

By interpolation of the above[7], the fractional norm estimate:

$$\|\mathcal{B}v\|_{1/2,\partial\Omega} \leq ch^{k}\|v\|_{1,\Omega}\,, \tag{4.46}$$

is also valid. If additionally $v \in H^2(\Omega)^4$, then the norms $\|v\|_{1,\Omega}$ in (4.44)-(4.46) can alternatively be replaced by $h^{1/2}\|v\|_{2,\Omega}$.

Since elements v of the trial subspace V_h do not necessarily satisfy the boundary condition $\mathcal{B}v = 0$ on the whole boundary, the existence of a unique solution for (4.41) has to be verified. In fact, it will be proved only for small enough parameters h. Furthermore the following equivalence result holds, cf. Braess [44]:

Lemma 4.7. *The minimum condition* (4.41) *is equivalent to the equation:*

$$a(u_h - u, v) = 0\,, \quad v \in V_h\,, \tag{4.47}$$

$$\Leftrightarrow \quad a(u_h, v) = (f, \mathcal{L}v)_0\,, \quad v \in V_h\,, \tag{4.48}$$

where $a(\cdot,\cdot)$ is the bilinear form corresponding to the norm ϱ:

$$a(u, v) = (\mathcal{L}u, \mathcal{L}v)_0 + \Lambda u \Lambda v\,. \tag{4.49}$$

Proof. This can be seen by a standard argument [44]. According to (4.41) set $\mathcal{J}(v) := \varrho^2(v - u)$ and for $v, u_h \in V_h$, $t \in \mathbb{R}$:

$$\begin{aligned}\mathcal{J}(u_h + tv) &= \|\mathcal{L}u_h + t\mathcal{L}v - f)\|_0^2 + |\Lambda u_h + t\Lambda v - c)|^2 \\ &= \mathcal{J}(u_h) - 2t(\mathcal{L}u_h - f, \mathcal{L}v)_0 + t^2\|\mathcal{L}v\|_0^2 \\ &= \mathcal{J}(u_h) - 2t[a(u_h, v) - (f, \mathcal{L}v)_0] + t^2\|\mathcal{L}v\|_0^2\,,\end{aligned} \tag{4.50}$$

holds. If the Galerkin orthogonality (4.48) is valid, (4.50) gives:

$$\mathcal{J}(u_h + tv) = \mathcal{J}(u_h) + t^2\|\mathcal{L}v\|_0^2 > \mathcal{J}(u_h)\,, \tag{4.51}$$

and therefore u_h describes a minimum of the functional $\mathcal{J}$. On the other hand, if u_h is a

[7] cf. Proposition B.14 with $m = 1$, $k = 1/2$ which is also valid for fractional-order trace norms.

minimum, then:

$$\frac{\mathrm{d}}{\mathrm{d}t} J(u_h + tv)\Big|_{t=0} \stackrel{!}{=} 0\,,$$

and therefore (4.48) holds. □

(b) Existence and uniqueness of the solution

In the case of a *conforming* FE discretisation one tries to approximate the solution u_h of (4.40) in a subspace of:

$$V = \left\{ v \in H^1(\Omega)^4 \;\middle|\; \mathcal{B}v = 0 \text{ on } \partial\Omega \right\}, \tag{4.52}$$

that is $u_h \in V_h \subset V$. However, as mentioned before, any kind of triangulation or domain discretisation in general will approximate the boundary only inexactly, i.e. the boundary conditions will be fulfilled only point-wise. Thus, a common FE discretisation involving Dirichlet boundary conditions is *non-conforming* in the sense that $V_h \not\subset V$ but at least $V_h \subset H^1(\Omega)^4$. In the conforming case an existence and uniqueness result for problem (4.41), or alternatively (4.47), is easily obtained by the classical theorem of Lax-Milgram, see [44], which requires the bilinear form to be elliptic. It will be shown that this also works in the present case; the relevant definitions are introduced below.

Definition 4.8. (Elliptic bilinear form) *Let $\tilde{V} \subset H^m$ be a Hilbert space. A bilinear form $a : \tilde{V} \times \tilde{V} \to \mathbb{R}$ is said to be* continuous *if a constant $\beta > 0$ exists such that:*

$$|a(u,v)| \leq \beta \|u\|_m \|v\|_m \quad \textit{for all } u, v \in \tilde{V}\,.$$

*A symmetric and continuous bilinear form $a(\cdot,\cdot)$ is said to be $\tilde{V}$-*elliptic*,* elliptic *or* coercive*, if another constant $\alpha > 0$ exists, such that:*

$$a(v,v) \geq \alpha \|v\|_m^2 \quad \textit{for all } v \in \tilde{V}\,.$$

Theorem 4.9. (Lax-Milgram) *Let $\tilde{V}$ be a Hilbert space and $a : \tilde{V} \times \tilde{V} \to \mathbb{R}$ a symmetric, V-elliptic bilinear form. Furthermore, let $\ell \in \tilde{V}'$ be a linear continuous functional such that:*

$$|\ell(v)| \leq c \|v\|_{\tilde{V}} \quad \textit{for all } v \in \tilde{V}\,,$$

for a constant $c > 0$. Then the variational equation:

$$a(u,v) = \ell(v) \quad \textit{for all } v \in \tilde{V}\,. \tag{4.53}$$

has exactly one solution $u \in \tilde{V}$.

The above theorem is sufficient to ensure the unique solution of (4.47), if the ellipticity requirements can be shown. This is summarised and proven for the case $\tilde{V} := V_h \subset H^1(\Omega)^4$ in the following Corollary, cf. [215, 251].

Corollary 4.10. (Existence and uniqueness) *If h is small enough, problem* (4.47) *with $u \in V$ according to* (4.52) *has a unique solution $u_h \in V_h \subset H^1(\Omega)^4$ which coincides with* (4.41)*. Furthermore the bilinear form* (4.49) *is V_h-elliptic and thus defines an inner product over V_h; the induced norm $\|\cdot\|_a = \sqrt{a(\cdot,\cdot)}$ is equivalent to $\|\cdot\|_1$.*

Proof. Clearly, by the Cauchy-Schwarz inequality and the already demonstrated $H^1(\Omega)$-continuity of λ, see (4.34), the bilinear form (4.49) is continuous since for any $u, v \in H^1(\Omega)^4$:

$$\begin{aligned} |a(u,v)| &= |(\mathcal{L}u, \mathcal{L}v)_0 + \Lambda u \Lambda v| \\ &\leq \|\mathcal{L}u\|_0 \|\mathcal{L}v\|_0 + \|\Lambda u\|_0 \|\Lambda v\|_0 \\ &\leq \beta(|u|_1|v|_1 + \|u\|_1\|v\|_1) \leq \beta\|u\|_1\|v\|_1 \,. \end{aligned}$$

Combining the inequality (4.46) with the ADN estimate (4.36) for $q = 0$ yields that for sufficiently small $h > 0$ the bilinear form $a(\cdot,\cdot)$ is also coercive. Firstly, the ADN estimate gives:

$$\|v\|_1 \leq c \left(\|\mathcal{L}v\|_0 + \|\mathcal{B}v\|_{1/2,\partial\Omega} + |\Lambda v| \right) ,$$

for all $v \in V_h$. With (4.46) the boundary term can be estimated as:

$$\|\mathcal{B}v\|_{1/2,\partial\Omega} \leq ch^k \|v\|_1 \leq ch^k \left(\|\mathcal{L}v\|_0 + \|\mathcal{B}v\|_{1/2,\partial\Omega} + |\Lambda v| \right) ,$$

with $k \geq 1$ and therefore:

$$\|\mathcal{B}v\|_{1/2,\partial\Omega} \leq \frac{1}{1 - ch^k} \left(\|\mathcal{L}v\|_0 + |\Lambda v| \right) ,$$

in which the denominator term on the right-hand side remains positive if h is sufficiently small, i.e. $h < c^{-1/k}$. This leads to:

$$\alpha \|v\|_1^2 \leq \|\mathcal{L}v\|_0^2 + |\Lambda v|^2 = a(v,v) \,, \tag{4.54}$$

for all $v \in V_h$ and a constant $\alpha > 0$. As a consequence of continuity and coerciveness, $a(\cdot,\cdot)$ defines an inner product over V_h and induces a norm $\|\cdot\|_a = \sqrt{a(\cdot,\cdot)}$ which is equivalent to the norm of the respective Hilbert space $\|\cdot\|_1$; i.e., for all $v \in V_h$ it is:

$$\sqrt{\alpha}\|v\|_1 \leq \|v\|_a \leq \sqrt{\beta}\|v\|_1 \,,$$

involving the two constants α, β of Definition 4.8. Finally it is easy to verify the continuity of the linear functional $\ell(v) := (f, \mathcal{L}v)_0$; for all $v \in V_h$ it is:

$$\ell(v) := (f, \mathcal{L}v)_0 \leq c\|f\|_0\|\mathcal{L}v\|_0 \leq c\|v\|_1 \,.$$

At this point, the required uniqueness result is obtained directly via the Theorem 4.9 of Lax-Milgram with $\tilde{V}$ set to V_h. □

4.3.3 Error of the least-squares method

(a) $H^1(\Omega)$-convergence of the method

Obviously the convergence rate of method (4.41), or alternatively (4.48), depends on the type of FE spaces, i.e. on the precise definition of the function space V_h in (4.42), and a necessary requirement for any reasonable proposition to state is knowledge associated with the principle approximation properties of the FE spaces under consideration. Here considerations are restricted to classical continuous and piece-wise polynomial Lagrange basis functions. According to [44], for each $u \in H^m(\Omega)^4$, $m \geq 2$ a unique interpolation function $v = I_h(u) \in V_h$ exists in Lagrange FE spaces and the corresponding interpolation error can be estimated. For elements $v \in V_h \subset H^1(\Omega)^4$ higher order Sobolev norms are not declared so that frequently grid-dependent norms are used as an alternative:

$$\|v\|_{m,h} := \left(\sum_{T_j \in \mathcal{T}_h} \|v\|^2_{m,T_j} \right)^{1/2} ,$$

which also occur in the following theorem related to interpolation error [44]. For the case $m = 1$ both norms are of the same magnitude, i.e. $\|v\|_1 \eqsim \|v_{1,h}\|$. The theorem relates to triangular elements, although similar results are valid and well-known [44] for other element types.

Theorem 4.11. (Interpolation error) *Let $k \geq 1$ and $\mathcal{T}_h$ be a quasi-uniform triangulation of Ω, then interpolation by piecewise polynomial trial functions of degree k gives:*

$$\|v - I_h v\|_{m,h} \leq c h^{k+1-m} |v|_{k+1,\Omega} \quad \textit{for } v \in H^{k+1}(\Omega)\,, \quad 0 \leq m \leq k+1\,, \tag{4.55}$$

for a constant $c = c(\Omega, \kappa, k)$.

The error of the least-squares FEM measured in Hilbert norms $\|\cdot\|_{m,h}$ must be greater or equal to the error of the best possible approximation in the associated approximation space V_h, indicated by (4.55); in this sense the inequality (4.55) defines a *lower error bound* so that a method with an error of the same order as that of $I_h v$ is considered *optimal*. In the following analysis optimal error estimates for the least-squares method are established

for different norms in which those estimates that are most interesting require the lowest regularity, that is $k = 1$, $m = 0, 1$. From these two cases, the estimate for $m = 1$ is by far easier to show: For instance in the conforming case $V_h \subset V$ where the boundary conditions are fulfilled exactly the minimum characterisation (4.41) would, by $H^1(\Omega)$-continuity of $a(\cdot, \cdot)$ and Λ, directly lead to the desired result. However, in the non-conforming case the boundary error has to be taken into account explicitly necessitating the boundary estimates from Assumption 4.6; this is shown along the lines of [215]:

Corollary 4.12. ($H^1(\Omega)$-convergence) *Let $u \in H^{k+1}(\Omega)^4$, $k \geq 1$ be a solution of problem (4.40) with homogeneous boundary conditions and let $\Lambda : H^1(\Omega) \to \mathbb{R}^\alpha$ be continuous. Then for the approximate solution $u_h \in V_h$ defined by (4.41) and using k^{th}-order Lagrange elements with respect to a given domain decomposition $\mathcal{T}_h$ the error estimate is:*

$$\|u - u_h\|_1 \leq c h^k \|u\|_{k+1} \,. \tag{4.56}$$

Proof. Let $I_h u \in V_h$ denote the piece-wise k^{th}-order polynomial interpolate of u, then the minimum characterisation (4.41) yields by $H^1(\Omega)$-continuity (4.34) of Λ:

$$\begin{aligned} \varrho(u - u_h) \leq \varrho(u - I_h u) &= \left(\|\mathcal{L}(u - I_h u)\|_0^2 + |\Lambda(u - I_h u)|^2 \right)^{1/2} \,, \\ &\leq c \|u - I_h u\|_1 \leq c h^k \|u\|_{k+1} \,. \end{aligned} \tag{4.57}$$

Alternatively, from the ADN-estimate (4.36) and the boundary estimate (4.46) it follows:

$$\|u - u_h\|_1 \leq \tilde{c} \left(\|\mathcal{L}(u - u_h)\|_0 + \|\mathcal{B}(u - u_h)\|_{1/2,\partial\Omega} + |\Lambda(u - u_h)| \right) \,, \tag{4.58}$$

$$\leq 2\tilde{c} \left(\varrho(u - u_h) + \|\mathcal{B} u_h\|_{1/2,\partial\Omega} \right) \,, \tag{4.59}$$

$$\leq c \left(h^k \|u\|_{k+1} + h^k \|u_h\|_1 \right) \,, \tag{4.60}$$

$$\leq c h^k \Big(\|u\|_{k+1} + \|u - u_h\|_1 + \|u\|_1 \Big) \,, \tag{4.61}$$

and therefore:

$$\begin{aligned} \|u - u_h\|_1 &\leq \frac{c}{1 - c h^k} h^k (\|u\|_{k+1} + \|u\|_1) \,, \\ &\leq \frac{2c}{1 - c h^k} h^k \|u\|_{k+1} \,, \end{aligned}$$

which clearly implies (4.56) if h is small enough, that is $h \leq c^{-1/k}$. This means, in order to preserve the convergence rate for higher order Lagrange elements the boundary has to be approximated to an appropriate higher order, too. This condition is fulfilled by the isoparametric FE approach, Sec. 5.1.3, where, for instance, linear elements are associated with a linear polygonal boundary approximation and quadratic elements with a piecewise quadratic boundary approximation. □

As a point of clarification it is explicitly emphasised that the error estimate (4.56) and all related h-dependent estimates rely on the additional regularity assumption that at least $u \in H^2(\Omega)$, which enters via interpolation Theorem 4.11. Theorem 4.2 provides a sufficient but extremely restrictive condition for $H^2(\Omega)$-regularity, namely that Ω is of class C^2 and $f \in H^1(\Omega)$. However, similar results based on weaker requirements for the boundary smoothness will be considered in Sec. 4.3.4.

In the case that the solution of system (4.35) and therefore the solution of the exact least-squares formulation is in $H^1(\Omega)$ only, inequality (4.56) is not valid any more but at least (4.55) can be replaced by the minimal interpolation statement:

$$\lim_{h\to 0} \|v - I_h v\|_1 = 0 \quad \text{for } v \in H^1(\Omega)\,,$$

which guarantees that (4.59) goes to zero and basic convergence of the least-squares approximation is achieved; thus for $u \in H^1(\Omega)$ only ($k = 0$), estimate (4.56) may be replaced by:

$$\lim_{h\to 0} \|u - u_h\|_1 = 0\,; \tag{4.62}$$

see Hackbusch [118], Proposition 8.4.6, p. 136, for the detailed underpinning argument in the conforming case.

(b) $L^2(\Omega)$-convergence of the method

According to Theorem 4.11 the interpolation error in the $L^2(\Omega)$-norm is one order better than in the $H^1(\Omega)$-norm which does not necessarily have to be the case for the error of the LSFEM. In order to demonstrate that the optimal $L^2(\Omega)$-estimate is indeed valid a duality argument is employed, often called the Nitsche trick [119]. In the present case this technique is applied by solving a second-order equation arising from the bilinear form $a(\cdot,\cdot)$ in a natural way, see Saranent and Wendland [215]. Consider the problem: given $f \in L^2(\Omega)^4$, find $w \in V$ such that:

$$a(w, v) = (f, v)_0 \quad \text{for all } v \in V\,, \tag{4.63}$$

involving the closed subspace $V \subset H^1(\Omega)^4$ according to (4.52). By ADN estimate (4.36) problem (4.63) has a unique solution $w \in V$ for all $f \in L^2(\Omega)^4$. The regularity of the solution w requires investigation which is carried out in detail in Appendix B.3 mainly providing two results. First, problem (4.63) has a unique solution $w \in H^2(\Omega)^4 \subset V$ satisfying the estimate:

$$\|w\|_2 \le c\|f\|_0\,. \tag{4.64}$$

Second, the Galerkin approximation $w_h \in V_h$ to problem (4.63), which is given by $a(w_h, v) = (f, v)_0$ for all $v \in V_h$, allows for the convergence result:

$$\|w - w_h\|_1 \leq ch\|w\|_2 \,. \tag{4.65}$$

These results provide the basis for the following optimal L^2-error estimate, the proof of which is close to [215]:

Theorem 4.13. *Let $u \in H^{k+1}(\Omega)^4$, $k \geq 1$ be a solution of problem (4.40) and let $\Lambda : H^1(\Omega)^4 \to \mathbb{R}^\alpha$ be continuous. Then for the approximate solution $u_h \in V_h$ defined by (4.41) and using k-th order Lagrange elements the error estimate is:*

$$\|u - u_h\|_0 + h\|u - u_h\|_1 \leq ch^{k+1}\|u\|_{k+1} \,. \tag{4.66}$$

Proof. Use $w \in V$ to denote the solution of:

$$a(w, v) = (e, v)_0 \quad \text{for all } v \in V \,,$$

with $e = u - u_h$. Then $e \in H^1(\Omega)^4$ holds but $e \notin V$ and therefore (B.50) gives:

$$\|e\|_0^2 = (e, e)_0 = a(w, e) - (\mathcal{C}'\mathcal{L}w, \mathcal{B}e)_{0,\partial\Omega} \,. \tag{4.67}$$

Next both terms on the right-hand side of equation (4.67) are estimated separately. According to the Galerkin orthogonality (4.47), for any $w_h \in V_h$ the following relation:

$$a(w_h, e) = a(w_h, u - u_h) = 0 \,,$$

is true, so that $a(w, e) = a(w - w_h, e)$. Thus the $H^1(\Omega)$-convergence (4.56) of u_h in combination with the similar result for the Galerkin approximation (4.65) yields the estimate:

$$\begin{aligned} a(w, e) = a(w - w_h, e) &\leq \|w - w_h\|_1 \|e\|_1 \\ &\leq ch\|w\|_2 h^k \|u\|_{k+1} \leq ch^{k+1}\|e\|_0 \|u\|_{k+1} \,, \end{aligned} \tag{4.68}$$

in which the last inequality is obtained via the regularity result (4.64) with $f := e$. Furthermore, the boundary term on the right-hand side of (4.67) is estimated as follows:

$$\begin{aligned} |(\mathcal{C}'\mathcal{L}w, \mathcal{B}u_h)_{0,\partial\Omega}| &\leq \|w\|_{1,\partial\Omega} \|\mathcal{B}(u - e)\|_{0,\partial\Omega} \\ &\leq c\|w\|_2 h^{(2k+1)/2}(\|e\|_1 + h^{1/2}\|u\|_2) \\ &\leq \tilde{c}\|e\|_0 h^{(2k+1)/2}(h^k\|u\|_{k+1} + h^{1/2}\|u\|_{k+1}) \\ &\leq \tilde{c}h^{k+1}\|e\|_0 \|u\|_{k+1} \,, \end{aligned} \tag{4.69}$$

using Assumption 4.6, particularly the last remark, from the first to the second line and

again the regularity result (4.64) from the second to the third line. Finally (4.67) together with the estimates (4.68) and (4.69) ensures the desired $L^2(\Omega)$-convergence:

$$\|u - u_h\|_0 \leq ch^{k+1}\|u\|_{k+1}\,.$$

and thus also (4.66). □

(c) Convergence of the non-homogeneous problem

Finally it is noted that the approximation of solutions for the complete problem (4.35) can be reduced to the approximation of a solution satisfying the homogeneous boundary condition [215]. Since the $m \times n$ matrix $\mathcal{B}$ can be extended to an invertible smooth matrix on $\partial\Omega$, it has a smooth right inverse $\mathcal{B}_r^{-1}$ on $\partial\Omega$. By continuing this $n \times m$ matrix smoothly into the interior of Ω and denoting a continuous right inverse of the mapping $\gamma_0 : u \to u|_{\partial\Omega}$, see Lions and Magenes [163], by $\gamma_{0,r}^{-1} : H^{k+1/2}(\partial\Omega)^m \to H^{k+1}(\Omega)^m$, a continuous right inverse $\tilde{\mathcal{B}}_r^{-1} : H^{k+1/2}(\partial\Omega)^m \to H^{k+1}(\partial\Omega)^n$ of $\mathcal{B}$ is defined by $\tilde{\mathcal{B}}_r^{-1}g = \mathcal{B}_r^{-1}\gamma_{0,r}^{-1}g$. Then, if u solves the boundary value problem:

$$\begin{cases} \mathcal{L}u = f \in H^k(\Omega)^n\,, \\ \mathcal{B}u = g \in H^{k+1/2}(\partial\Omega)^m\,, \\ \Lambda u = c\,, \end{cases} \tag{4.70}$$

take $w := \tilde{\mathcal{B}}_r^{-1}g$ and approximate the solution $v \in H^{k+1}(\Omega)^n$ of the system:

$$\begin{cases} \mathcal{L}v = f - \mathcal{L}w \in H^k(\Omega)^n\,, \\ \mathcal{B}v = 0\,, \\ \Lambda v = c - \Lambda w\,, \end{cases}$$

by means of an element $v_h \in V_h$ as described above. For the approximate solution:

$$u_h = w + v_h\,, \tag{4.71}$$

the following more general form of Theorem 5.2 from [215] holds:

Theorem 4.14. *Let $u \in H^{k+1}(\Omega)$, $k \geq 1$ be a solution of (4.70) and let u_h be the approximation given by (4.71) using k-th order nodal Lagrange elements. If Λ is $L^2(\Omega)$-continuous, the associated error estimate is:*

$$\|u - u_h\|_0 + h\|u - u_h\|_1 \leq ch^{k+1}(\|f\|_k + \|g\|_{k+1/2,\partial\Omega} + |c|)\,.$$

4.3.4 Least-squares FEM in irregular domains

Several elliptic boundary value problems have the fortunate property of a guaranteed smooth solution as long as the data and domain are smooth. However, many problems of interest are posed in non-smooth domains and, as a consequence, lose this property at a finite number of points on the boundary. In this section problems that have non-smooth solutions at "irregular boundary points", that is, points that are corners of polygonal domains, locations of changing boundary condition type, or both, are considered.

Standard solution techniques applied to such boundary value problems suffer from a global loss of accuracy due to the reduced smoothness of the solution [33, 243]. Several approaches are used to combat this so-called "pollution effect". The most common of these is systematic local mesh refinement near the singularities [12, 18] in which approximate knowledge of the singular behaviour may be crucial to design the precise refinement rates and to restore optimal convergence rates [164]. But even local refinement of FE subspaces of H^1 fails to converge to a solution that is not in H^1. If a basis for the singular functions is explicitly known, they can be incorporated directly into the FE space, sometimes called the "singularity subtraction method" [33, 57, 243]. For general 2D elliptic problems of a Poisson type and the plane elasticity and Stokes equations, the singular basis functions are known, see [149, 151] and [114, 115], respectively. There exist quite general approaches for plane ADN-elliptic systems [72] while for other problems, or in 3D, the character of the singular functions is less well understood.

In particular, the accuracy of least-squares methods in the presence of corner singularities has been investigated in [75]. In contrast to mixed Galerkin FEs, in the first-order least-squares setting it is much more natural to tackle the reduced convergence near singular points by adapting the *least-squares functional* instead of the *mesh*; this opens up the possibility of restoring convergence simply by minimising an appropriate *weighted* functional, depending on the distance to singular points, and to save computational cost induced by local mesh refinement. Such an approach was suggested for instance by Lee et al. [157] for 2D div-curl systems which is also related to the idea of energy-correcting methods presented in [136].

The simplest method of restoring convergence near irregular points is considered first, namely an appropriate mesh refinement. Consider a finite sector:

$$Q = \{(r, \varphi) : 0 < r < R, 0 < \varphi < \theta < 2\pi\}, \tag{4.72}$$

with corresponding edges:

$$S_0 = \{(r, \varphi) : 0 < r < R, \varphi = 0\}, \tag{4.73}$$

$$S_1 = \{(r, \varphi) : 0 < r < R, \varphi = \theta\}, \tag{4.74}$$

which serves as a localised model of a polygonal domain in the vicinity of a corner. Solution singularities for 2D elliptic systems can be described locally as functions $f(r,\varphi) = r^a$ in polar coordinates, in which the exponent a may depend on the corner angle as well as on the boundary conditions on both sides of the corner; such a singularity is placed at the origin of the sector (4.72). A simple calculation demonstrates that $f \in H^k(Q)$ only if $k < a+1$ which is a necessary condition for the k-th derivative of f to remain finite in the $L^2(Q)$-norm, i.e. $\|f^{(k)}\|_0 < \infty$:

$$\begin{aligned}\|f^{(k)}\|_0^2 &= \int_0^\theta \int_0^R \left| \frac{a!}{(a-m)!} r^{a-k} \right|^2 r \mathrm{d}r\mathrm{d}\varphi \\ &= \frac{\theta (a!)^2}{(a-k)!(a-k+1)!} \left(R^{2(a-k+1)} - \lim_{r\to 0} r^{2(a-k+1)} \right) .\end{aligned}$$

In order to guarantee the $H^2(Q)$-regularity of the qualitative h-dependent estimate of Theorem 4.14, requires $a > 1$ and for a minimal convergence result, cf. Lemma 4.4, at least $a > 0$. Thus it is of great importance to identify the possible radial exponents occurring in the corner solution of the first integral equations. Consider for now the reduced auxiliary problem:

$$\begin{cases} \mathcal{L}u = f \quad \text{in } Q\,, \\ \mathcal{B}u = [g^0, g^1] \quad \text{on } \partial Q\,, \\ \Lambda u = 0\,, \end{cases} \tag{4.75}$$

with u being defined by the convention (4.16) and involving the first integral Stokes operator from (4.15), the necessary constraints Λ, and boundary conditions $\mathcal{B}u = [g^0, g^1]$ with:

$$\mathcal{B}u|_{\Gamma_0} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} v = g^0 \quad \text{and} \quad \mathcal{B}u|_{\Gamma_1} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} v = g^1\,. \tag{4.76}$$

The complete sector boundary is given by $\partial Q = S_0 \cup S_1 = \Gamma_0 \cup \Gamma_1$ and three cases are distinguished: (i) $\Gamma_0 = S_0 \cup S_1$, $\Gamma_1 = \emptyset$, (ii) $\Gamma_0 = \emptyset$, $\Gamma_1 = S_0 \cup S_1$, and (iii) $\Gamma_0 = S_0$, $\Gamma_1 = S_1$; these correspond to pure velocity (i) or pure potential (ii) conditions on both edges as well as mixed conditions (iii). The latter is the most complicated, involving a corner point which at the same time is a meeting point of two different boundary conditions.

Analogously to [114, 115, 136, 157] the solution of the first integral equations in domains containing an irregular point is conveniently described in terms of weighted Sobolev spaces H^k_β of Kondrat'ev type [149] in which the weighting factor depends on the corner singularity. The space H^k_β comprises all distributions v for which the weighted Sobolev norm:

$$\|v\|^2_{k,\beta} = \sum_{|j|\le k} \|r^{\beta-k+j} D^j v\|_0^2\,, \tag{4.77}$$

remains finite and forms a Hilbert space; the corresponding semi-norm with $|j| = k$ is denoted by $|v|_{k,\beta}$. For the Kondrat'ev spaces the embeddings $H^{m+1}_\beta \subset H^m_{\beta-1} \subset H^m_\beta$ and $H^m_0 \subset H^m$ are valid for $m \in \mathbb{N}$, $\beta \in \mathbb{R}$. In terms of these function spaces, the following result is derived for an infinite sector Q with $R \to \infty$ in detail in Appendix B.4:

Theorem 4.15. *Let $f \in H^k_\beta(Q)^4$, $g^l \in H^{k+\frac{1}{2}}_\beta(\Gamma^l)^2$, $l = 0, 1$ with $k \geq 0$. Then the sector problem (4.75), (4.76) has a unique solution $u \in H^{k+1}_\beta(Q)^4$, provided that $k-\kappa < \beta < k+\kappa$, and the following a priori estimate holds:*

$$\|u\|_{H^{k+1}_\beta(Q)} \leq C \left(\|f\|_{H^k_\beta(Q)} + \sum_{l=0,1} \|g^l\|_{H^{k+1/2}_\beta(\Gamma^l)} \right) . \tag{4.78}$$

The value of κ which effectively determines the range of β is given by:

$$1 + \kappa = \min_{\mathrm{Re}(z)>1} \mathrm{Re}(z) = \min_{\mathrm{Re}(z)\neq 1} |\,\mathrm{Re}(z)|\,, \tag{4.79}$$

and depends on the precise boundary conditions and the sector angle θ:

1. *For either pure velocity boundary conditions (i) or pure stress boundary conditions (ii), the z-values in (4.79) are solutions of:*

$$\sin^2((z-1)\theta) = (z-1)^2 \sin^2(\theta)\,; \tag{4.80}$$

2. *for mixed boundary conditions (iii), the z-values satisfy:*

$$\cos^2((z-1)\theta) = (z-1)^2 \sin^2(\theta)\,. \tag{4.81}$$

Firstly note, that the above theorem provides a suitable function space setting, i.e. weighted Sobolev spaces, in terms of which the data and solution of an irregular boundary value problem can be described. Inequality (4.78) provides a corresponding regularity shift which in contrast to the ADN-result (4.36) captures the case of such an irregular boundary point. This is helpful because it guarantees smoother solutions in irregular domains if only the data f and g is smooth enough on the regular parts of the domain boundary which is otherwise not obvious. Definition (4.79) demonstrates that the β-weights are effectively determined by the real part $\mathrm{Re}(z)$ of solutions z of two transcendental equations (4.80) and (4.81); namely determined by the critical solution z_κ which is closest to the line $\mathrm{Re}(z) = 1$. Fig. 4.1 visualises the solutions versus the sector angle θ and shows an axis-symmetry with respect to $\mathrm{Re}(z) = 1$.

When using a H^1 FEM, a natural case to consider is $k = 1$. If for instance $\mathrm{Re}(z_\kappa) > 2+\varepsilon$, it follows $\kappa > 1+\varepsilon$ and $u \in H^2_\beta(Q)$ for all $-\varepsilon < \beta < 2+\varepsilon$; in particular the finiteness of the weighted Sobolev norm (4.77) for $\beta = 0, 1, 2$ implies the standard regularity $u \in H^2(Q)$.

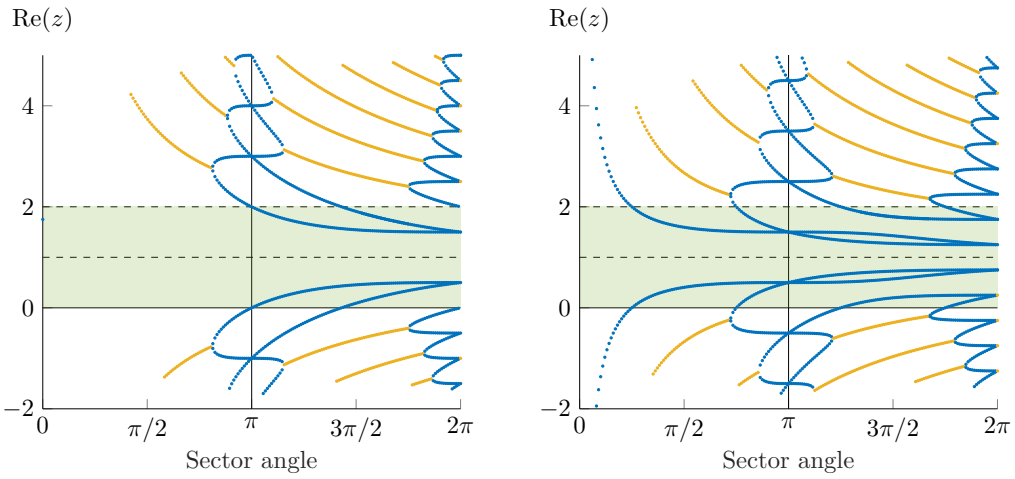


Figure 4.1: Real part of the solutions z of the transcendental equation (4.80) (left) and (4.81) (right) versus the sector angle θ. The blue dots correspond to z-values with non-zero imaginary part while yellow dots correspond to those with zero imaginary part. The green corridor belongs to solutions which are not in $H^2(Q)$.

In this respect, according to Fig. 4.1 (left), a corner with pure Dirichlet velocity or pure Dirichlet potential boundary conditions will not affect the optimal H^1-convergence as long as $\theta < \pi$ which corresponds to the classical result for Stokes flow in primitive variables and convex domains, see [118] Proposition 12.2.19. However, for concave domains $\theta > \pi$ the solution $\mathrm{Re}(z_\kappa)$ is in the critical green corridor and the regularity falls below H^2; even more severe, for mixed boundary conditions this critical point is already reached for approximately $\theta > \pi/4$. Additionally, from Fig. 4.1 it can be seen that the low regularity is caused by a small number of singular solutions which grows with the sector angle, reaching up to two in the left-hand case and up to four in the right-hand case; critical angles can be identified for the different (blue) "solution lines" entering the green corridor.

In order to illustrate the consequences of the above result in more detail, consider how the proof of optimal H^1-convergence, Corollary 4.12, is influenced for corner domains with critical angles; for convenience assume that the boundary is captured exactly, so that only the first part of the proof is relevant. Furthermore, an interpolation result in weighted Sobolev spaces, similar to Theorem 4.11, is needed [136, 157]:

Theorem 4.16. (Interpolation error) *Let $k > 0$, $\beta - 1 \leq \alpha \leq \beta$ and $\mathcal{T}_h$ a quasi-uniform triangulation of Ω, then the interpolation by piecewise polynomial trial functions of degree k leads to:*

$$\|v - I_h v\|_{m,\alpha,h} \leq c h^{k+1-m+\alpha-\beta} \|v\|_{k+1,\beta} \quad \text{for } v \in H^{k+1}_\beta(\Omega)\,, \quad 0 \leq m \leq k+1\,, \tag{4.82}$$

with a constant $c = c(\Omega, \beta, k)$.

Now, consider the relevant case that under certain assumptions on the data a solution in a domain Ω with a single irregular point is locally, i.e. near that point, identified as $u \in H^2_\beta(\Omega) \cap H^1(\Omega)$ but $u \notin H^2(\Omega)$; for instance $\mathrm{Re}(z_\kappa) = 1 + \varepsilon$ and $1 - \varepsilon < \beta < 1 + \varepsilon$. Then by adopting estimate (4.57) as well as subsequent estimates, together with use of (4.82) with $m = k = 1$ and $\alpha = 0$, the following result is obtained:

$$\begin{aligned} \|u - u^h\|_1 &\leq c\|\mathcal{L}(u - u^h)\|_0 \leq c\|\mathcal{L}(u - I_h u)\|_0 \leq c\|u - I_h u\|_1 \\ &\leq c\|u - I_h u\|_{1,\alpha=0} \leq ch^{1-\beta}\|u\|_{2,\beta}\,, \end{aligned} \tag{4.83}$$

which reflects the drop in convergence from a rate of h to $h^{1-\beta}$, depending on the corner-related parameter β. An obvious remedy to restore optimal convergence is systematic mesh refinement near the irregular point in the following form. Assuming that far away from the corner a regular and approximately homogeneous triangulation is established with a constant mesh width parameter h, then for triangles T approaching the corner the circumradius h_T should decrease according to the rule:

$$h_T \leq h^{\frac{1}{1-\beta}}\,. \tag{4.84}$$

Such a rule can easily be adopted for domains with several corners and for each corner local mesh refinement performed with a different angle-dependent β.

However, a generalisation of the above kind would require a proper extension of Theorem 4.15 to more general domains with corners. Consider, for instance, an arbitrary smooth domain Ω with, lets say, a single irregular point and for simplicity assume that the irregular point is again located at the coordinate origin so that a suitably truncated sector Q is part of the domain Ω. The procedure can in principle be adopted from [114] so only the main idea is sketched out below.

A cut-off function $\phi_\delta(r) \in C^\infty(\mathbb{R})$ can be defined such that $\phi_\delta = 1$ for $0 < r < \delta/2$ and $\phi_\delta = 0$ for $r > \delta$. Then, a suitable solution $u_Q \in H^2_\beta(Q)$, $1 - \varepsilon < \beta < 1 + \varepsilon$ of the infinite sector problem according to Theorem 4.15 with $k = 1$ can be truncated as $u_\delta = \phi_\delta u_Q$ and shown to be locally identical to the solution $u \in H^1(\Omega)$ of the global problem. As a consequence it can be shown that even $u \in H^2_\beta(\Omega)$ in the whole of Ω and thus the shift inequality (4.78) is valid for a more general domain Ω as well. Although similar generalisations would be possible for $k > 1$, for reasons explained by Guo and Schwab [115] it is useful to consider the general shift theorem for $k > 1$ in a modified weighted Hilbert norm, which due to the restricted definition (4.77) is obviously also necessary for domains with multiple corners. Also note, that for increasing k the estimate $\|u - u^h\|_1 \leq ch^{k-\beta}\|u\|_{k+1,\beta}$ for k-th order Lagrange elements is less helpful because β grows with k and thus an unacceptable strong mesh refinement will be necessary to retain k-th order convergence.

As mentioned at the outset of this subsection, in the context of a LSFEM it could be more desirable to retain optimal convergence rates simply by modifying the norm in which the residual is minimised, that is to replace the $L^2(\Omega)$-norm in (4.41) by some $L^2_\beta(\Omega)$-norm as shown by Lee et al. [157]. Such a method would, in contrast to (4.83), allow for an optimal $H^1(\Omega)$ error estimate for k-th order Lagrange elements of the form:

$$\begin{aligned}\|u-u^h\|_{1,\beta} &\leq c\|\mathcal{L}(u-u^h)\|_{0,\beta} \leq c\|\mathcal{L}(u-I_h u)\|_{0,\beta} \leq c\|u-I_h u\|_{1,\beta} \\ &\leq ch^k\|u\|_{k+1,\beta}\,, \end{aligned} \tag{4.85}$$

if $u \in H^{k+1}_\beta(\Omega)$. The above is particularly useful when $u \in H^1 \setminus H^2$ but also opens up the treatment of even less regular or discontinuous problems, i.e. mixed boundary conditions with concave angles or the lid-driven cavity case with discontinuous boundary conditions, when the ideas in [157] are followed further. The latter reference also provides arguments how optimal L^2-convergence can be recovered from the above H^1-estimate. However, the validity of the first inequality in (4.85) remains unclear thus far when general irregular domains are considered (beyond the sector), because the above matching arguments, utilising the ADN-estimate (4.36), do not translate to the case $u \in H^1_\beta \not\subset H^1(\Omega)$ and thus a generalisation of (4.78) when $k = 0$ remains to be investigated.

Finally, it is mentioned that the above investigations for irregular domains are confined to Stokes flow but most of the results carry over to NS flow as well because the asymptotic solution near the boundary is strongly dominated by viscous effects while the convective terms can more or less be neglected [166].

4.4 Analysis of the non-linear problem

4.4.1 Non-linear variational problem

This section explores the convergence properties of the least-squares discretisation of problem (4.17) in order to clarify in which ways the results of Theorem 4.14 for the Stokes case extend to the non-linear case. Consider the NS solution $u + \hat{u}$ as the sum of the linear Stokes solution $\hat{u}$ of (4.70) and a perturbation field u. Obviously the following non-linear and homogeneous first-order system:

$$\mathcal{L}_{ij}u_j + \mathcal{N}_{ij}[u+\hat{u}](u_j+\hat{u}_j) = 0\,, \tag{4.86a}$$

$$\mathcal{B}u = 0\,, \tag{4.86b}$$

$$\Lambda u = 0\,, \tag{4.86c}$$

results from (4.17) when both the boundary conditions and the constraints are linear. Similar to (4.41), the above system (4.86) may be solved by minimisation of the $L^2(\Omega)$

least-squares functional:

$$\mathcal{J}(u) = \sum_{i=1}^{4} \|\mathcal{L}_{ij}u_j + \mathcal{N}_{ij}[u+\hat{u}](u_j+\hat{u}_j)\|_0^2 + \sum_{i=1}^{\alpha} |\Lambda_{ij}u_j|^2 \,, \tag{4.87}$$

in the function space V according to (4.52); this yields the minimisation task:

$$\begin{aligned} &\text{Find } u = (u_1, u_2, \phi_1, \phi_2) \in V, \text{ such that:} \\ &\quad J(u) \leq J(v) \quad \forall\, v \in V \,. \end{aligned} \tag{4.88}$$

Accordingly, the FE approximation in a suitable finite-dimensional subspace V_h of V:

$$V \supset V_h = \left\{ v \in H^1(\Omega)^4 \cap C^0(\overline{\Omega})^4 \,\middle|\, \mathcal{B}v = 0 \text{ on } \partial\Omega \right\} \,, \tag{4.89}$$

leads to the problem:

$$\begin{aligned} &\text{Find } u^h \in V_h, \text{ such that:} \\ &\quad J(u^h) \leq J(v) \quad \forall\, v \in V_h \,. \end{aligned} \tag{4.90}$$

Note, that in contrast to (4.42) the definition of the subspace V_h in (4.89) assumes a conforming FE discretisation in which the boundary is exactly captured by the triangulation and therefore boundary conditions are fulfilled exactly. This assumption, however, is made for convenience only. In principle the preceding analysis of the linear problem revealed that the error in the boundary discretisation may be neglected as long as the grid parameter h is small enough and the order of the boundary approximation is sufficient; a finding that certainly translates to the non-linear problem as well, so that unnecessary complications due to an inexact boundary approximation, which will not affect the principle results, are avoided.

Adapting the consideratons in the proof of Lemma 4.7 it is easy to verify that minimisation problem (4.88) is equivalent to the following variational setting:

$$\begin{aligned} &\text{Find } u = (u_1, u_2, \phi_1, \phi_2) \in V, \text{ such that } \forall\, v \in V : \\ &a(u,v) := (\mathcal{L}u + \mathcal{N}[u+\hat{u}](u+\hat{u}), \mathcal{L}v + 2\mathcal{N}[u+\hat{u}]v)_0 + \Lambda u \Lambda v = 0 \,, \end{aligned} \tag{4.91}$$

in which the commutation symmetry $\mathcal{N}[u]v = \mathcal{N}[v]u$ has been utilised. Then the least-squares discretisation method for the NS equations is defined either by (4.88) or (4.91) in which u, v and V may be replaced by u^h, v^h and V_h in order to indicate the approximate quantities. For the FE space V_h, again the approximation property (4.55) is assumed.

4.4.2 Discretisation error estimates

The main goal of this subsection is to derive error estimates for the non-linear least-squares problem (4.90), beginning by introducing the relevant function spaces:

$$V^m = H^{m+1}(\Omega)^4 \cap V\,, \quad Y = V^*\,, \quad Z = L^{3/2}(\Omega)^4\,, \tag{4.92}$$

for some non-negative integer m, in which V^* denotes the dual of V with respect to the $L^2(\Omega)$ inner product, see Appendix B. An important key feature for the treatment of the non-linear problem is the fact that system (4.91) can be recast in the canonical form:

$$F(\lambda, u) := u + T \cdot G(\lambda, u) = 0\,, \tag{4.93}$$

or equivalently in the discrete form:

$$F^h(\lambda, u) := u^h + T_h \cdot G(\lambda, u^h) = 0\,, \tag{4.94}$$

which proves extremely useful when applying the abstract approximation theory for non-linear PDEs developed by Brezzi et al. [54] in the 1980's. The theory is well summarised for the NS equations in the book of Girault and Raviart [107] and in the context of least-squares methods as applied by Bochev et al. [35]; the latter reference also proved an excellent guide for developing the present analysis.

The discrete form (4.94) is introduced by way of operator T_h and therefore, the error estimates will depend largely on the nature of T and its approximation T_h. Considering the original system (4.86), the representations (4.93) and (4.94) can obviously be achieved when T is identified to be the unique Stokes solution operator and its FE counterpart, that is $u = Tg$ for any $g \in Y$ if and only if $\mathcal{L}u = g$, $\mathcal{B}u = 0$ and $\Lambda u = 0$. If T is known, then operator G is defined by $G(\lambda, u) := \mathcal{N}[u + \hat{u}](u + \hat{u})$ in which the problem parameter λ is identified as the Reynolds number which enters via the definition of the operator $\mathcal{N}$ in (4.15). Subsequently, the Reynolds number is assumed to be taken from a fixed compact set $D_\lambda \subset \mathbb{R}^+$, i.e. $\lambda := \mathrm{Re} \in D_\lambda$. Thus, G is an operator $G : D_\lambda \times V \to Y$ so that the composition $TG(\lambda, u) : D_\lambda \times V \to V$ is well defined. However, in the remaining analysis the non-linear system (4.86) is not considered directly but rather its variational formulation (4.91) leading to differing weak definitions of the operators T and G, similar to those in [35], where for instance G involves a coupling of $\mathcal{L}$ and $\mathcal{N}$[8]:

[8]Note, that $(g, v) := g(v)$ denotes the duality pair, if $g \in V^*$ is interpreted as a linear functional $g : V \to \mathbb{R}$ from the dual space of V. Otherwise, g can be identified with an element of the function space $H^{-1}(\Omega)^4 \supset L^2(\Omega)^4$, see [118], so that $(g, v) := (g, v)_0$ can also be interpreted as the continuous extension of the $L^2(\Omega)$ scalar product $(\cdot, \cdot)_0$ from $L^2(\Omega)^4 \times L^2(\Omega)^4$ to $H^{-1}(\Omega)^4 \times H^1(\Omega)^4$.

Definition 4.17. *For $g \in Y$ the operator $T : Y \to V$ is defined by $u = Tg$ if and only if:*

$$a_s(u, v) := (\mathcal{L}u, \mathcal{L}v)_0 + \Lambda u \Lambda v = (g, v)\,, \quad \text{for all } v \in V\,. \tag{4.95}$$

Correspondingly, $T_h : Y \to V_h$ is defined by $u^h = T_h g$ for $g \in Y$ if and only if:

$$a_s(u^h, v^h) = (g, v^h)\,, \quad \text{for all } v^h \in V_h\,. \tag{4.96}$$

On the other hand, for $u \in V$ the operator $G : D_\lambda \times V \to Y$ is defined by $g = G(\lambda, u)$ if and only if:

$$(g, v) = (\mathcal{L}u, 2\mathcal{N}[u + \hat{u}](v))_0 + (\mathcal{N}[u + \hat{u}](u + \hat{u}), \mathcal{L}v + 2\mathcal{N}[u + \hat{u}](v))_0\,, \tag{4.97}$$

for all $v \in V$.

Obviously the following equivalences hold [35]:

Lemma 4.18. *Assume that T, T_h and G are defined by* (4.95), (4.96) *and* (4.97)*, respectively, then the non-linear problem* (4.91) *is equivalent to* (4.93) *and the corresponding discret non-linear problem is equivalent to* (4.94)*.*

Error estimates for the least-squares formulation (4.90) are now derived from abstract approximation theory [54, 107]. Below the main result of this theory for general T and T_h, but otherwise specialised to the needs of the present analysis, is taken from [35]. Let $D_u G(\lambda, u)$ and $D_u F(\lambda, u)$ denote the Fréchet derivatives of G and F with respect to u, see Appendix B.1.5, and consider the following [35]:

Definition 4.19. *The set $\{(\lambda, u(\lambda)) | \lambda \in D_\lambda\}$ is referred to as a* regular branch of solutions *of* (4.93) *if $u = u(\lambda)$ is a weak solution of* (4.93) *for each $\lambda \in D_\lambda$, if $\lambda \mapsto u(\lambda)$ is a continuous map $D_\lambda \to V$ and if $D_u F(\lambda, u)$ is an isomorphism of V.*

Equipped with these definitions the following theorem holds according to [35]:

Theorem 4.20. *Let $F(\lambda, u) = 0$ denote abstract form* (4.93) *and assume that $\{(\lambda, u(\lambda)) | \lambda \in \Lambda\}$ is a branch of regular solutions of* (4.93)*. Furthermore, assume that $T \in L(Y, V)$, that G is a C^2 map $D_\lambda \times V \to Y$ such that all second derivatives of G are bounded on bounded subsets of $D_\lambda \times V$ and that there exists a space $Z \subset Y$, with continuous embedding, such that $D_u G(\lambda, u) \in L(V, Z)$ for all $\lambda \in D_\lambda$ and $u \in V$. If the approximate problem* (4.94) *is such that:*

$$\lim_{h \to \infty} \|(T - T_h) g\|_V = 0\,, \tag{4.98}$$

for all $g \in Y$ and:

$$\lim_{h \to \infty} \|T - T_h\|_{L(Z,V)} = 0\,, \tag{4.99}$$

then:

(i) There exists a neighbourhood B of the origin in V and, for h sufficiently small, a unique C^2 function $\lambda \mapsto u^h(\lambda) \in V_h$ such that $\{(\lambda, u^h(\lambda)) | \lambda \in D_\lambda\}$ is a branch of regular solutions of the discrete problem (4.94) and $u(\lambda) - u^h(\lambda) \in B$ for all $\lambda \in D_\lambda$.

(ii) For all $\lambda \in D_\lambda$ the following holds:

$$\|u(\lambda) - u^h(\lambda)\|_V \leq c\|(T - T_h)G(\lambda, u(\lambda))\|_V \,. \tag{4.100}$$

(iii) If the regular branch is such that $u(\lambda) \in V^m$ for some integer $m \geq 1$ and $\tilde{d} := \min\{d, m\}$, where d is the largest integer satisfying the approximation property (4.55), then:

$$\|u(\lambda) - u^h(\lambda)\|_1 \leq ch^{\tilde{d}}\|u(\lambda)\|_{\tilde{d}+1} \,. \tag{4.101}$$

In the next few lemmas, the hypotheses of Theorem 4.20 are verified for the least-squares formulation. The starting point is that of establishing essential properties of the operators T and T_h, which are defined by (4.95), (4.96), respectively. The argument follows closely that of Bochev et al. [35].

Lemma 4.21. *$T \in L(Y, V)$ and $T_h \in L(Y, V_h)$.*

Proof. The bilinear form $a_s(\cdot, \cdot)$ is continuous and coercive, as already shown in the proof to Corollary 4.10 and if, for convenience only, the boundary approximation is assumed to be exact, then by virtue of the inclusion $V_h \subset V$, it is also continuous and coercive on $V_h \times V_h$. Furthermore, for each $g \in Y$, (g, v) defines a continuous functional on V, such that the Theorem 4.9 of Lax-Milgram is applicable. As a consequence, for all $g \in Y$, variational problems (4.95), (4.96) have unique solutions $u \in V$ and $u^h \in V_h$ respectively, i.e. $T : Y \to V$ and $T_h : Y \to V_h$ are well defined linear operators. From:

$$c\|u\|_V^2 \leq a_s(u, u) = (g, u) \leq \|g\|_Y \|u\|_V \,,$$

it follows that:

$$\|Tg\|_V = \|u\|_V \leq c\|g\|_Y \,,$$

i.e., T is in $L(Y, V)$. The proof that $T_h \in L(Y, V_h)$ is similar. □

Lemma 4.22. [35] *Convergence properties* (4.98) *and* (4.99) *hold. If, in addition,* $g \in Y$ *is such that* $Tg \in V^m$ *for some* $m \geq 1$ *and* $\tilde{d} = \min(d, m)$*, where* d *is the largest integer satisfying* (4.55)*, then:*

$$\|(T - T_h)g\|_V \leq ch^{\tilde{d}} \|Tg\|_{V^{\tilde{d}+1}} .$$

Proof. Firstly, the convergence statement (4.98) is established, i.e. that:

$$\|(T - T_h)g\|_V = \|u - u^h\|_1 \to 0 ,$$

when $h \to 0$. Recall that $T : Y \to V$ and therefore from $g \in Y$ it follows that $Tg = u \in V$ and $T_h g \in V_h \subset V$, i.e. $u, u_h \in H^1(\Omega)^4$. Then the above limit follows from the basic convergence result (4.62), cf. the considerations in Sec. 4.3.3(a).

Secondly, note that (4.99) follows from (4.98) when the embedding $Z \subset Y$ is compact; then the following estimate holds for $h \to 0$:

$$\|T - T_h\|_{L(Z,V)} := \sup_{0 \neq g \in Z} \frac{\|(T - T_h)g\|_V}{\|g\|_Z} \leq \sup_{0 \neq g \in Y} \frac{\|(T - T_h)g\|_V}{\|g\|_Y} \to 0 .$$

The compact embedding $Z \subset Y$ is valid because according to Theorem B.15, see Appendix B, $Y^* = V \subset H^1(\Omega)^4$ is compactly embedded in $L^3(\Omega)^4$; therefore $Z = L^{3/2}(\Omega)^4 = (L^3(\Omega)^4)^*$ is compactly embedded in Y.

To prove the last part of the lemma, suppose $u = Tg \in V^m$. Then an immediate consequence of the $H^1(\Omega)$ convergence result of Corollary 4.12 is the Stokes error estimate:

$$\|(T - T_h)g\|_V = \|u - u^h\|_1 \leq ch^{\tilde{d}} \|u\|_{\tilde{d}+1} .$$

□

The only hypotheses of Theorem 4.20 that remains to be verified are the assumptions concerning the non-linear operator G. For this purpose, the weak and strong forms of the first Fréchet derivative $D_u G(\lambda, u)$ and the weak form of the second Fréchet derivative $D_u^2 G(\lambda, u)$ are needed. Recall the characterisation (B.19) in Appendix B.1.5:

$$D_u G(\lambda, u)h = G(\lambda, u + h) - G(\lambda, u) + o(\|h\|_V) ,$$

for $h \in V$, which yields the following weak representation of $D_u G(\lambda, u)$ [35]:

Definition 4.23. *For given* $u \in V$*,* $\lambda \in D_\lambda$ *the mapping* $D_u G(\lambda, u) : V \to Y$*,* $h \mapsto D_u G(\lambda, u)h$ *is defined by* $g = D_u G(\lambda, u)h$ *if and only if for all* $v \in V$*:*

$$\begin{aligned}(g, v) = {} & (\mathcal{L}u + \mathcal{N}[u + \hat{u}](u + \hat{u}), 2\mathcal{N}[h]v)_0 + (\mathcal{L}h, 2\mathcal{N}[u + \hat{u}]v)_0 \\ & + (2\mathcal{N}[u + \hat{u}]h, \mathcal{L}v + 2\mathcal{N}[u + \hat{u}]v)_0 + o(\|h\|_V) ,\end{aligned} \tag{4.102}$$

where all $o(\|h\|_V)$-terms have been neglected. The strong form of $D_u G(\lambda, u)h$ can be found from (4.102) using standard integration by parts leading to the following abstract representation:

$$\begin{aligned} g/2 = \mathcal{N}^T[h]\mathcal{L}u + \mathcal{N}^T[h]\mathcal{N}[u+\hat{u}](u+\hat{u}) + \mathcal{N}^T[u+\hat{u}]\mathcal{L}h \\ + \mathcal{L}^*\mathcal{N}[u+\hat{u}]h + 2\mathcal{N}^T[u+\hat{u}]\mathcal{N}[u+\hat{u}]h + o(\|h\|_V)\,. \end{aligned} \tag{4.103}$$

Finally, the characterisation:

$$D_u G(\lambda, u)[h, \hat{h}] = G(\lambda, u+\hat{h})h - G(\lambda, u)h + o(\|h\|_V, \|\hat{h}\|_V)\,,$$

is employed for the second Fréchet derivative such that its weak form is given by [35]:

Definition 4.24. *For given $u \in V$, $\lambda \in D_\lambda$ the mapping $D_u^2 G(\lambda, u) : V \times V \to Y$, $(h, \hat{h}) \mapsto D_u^2 G(\lambda, u)[h, \hat{h}]$ is defined by $g = D_u^2 G(\lambda, u)[h, \hat{h}]$ if and only if for all $v \in V$:*

$$\begin{aligned} (g, v) = (\mathcal{L}\hat{h} + 2\mathcal{N}[u+\hat{u}]\hat{h}, 2\mathcal{N}[h]v)_0 + (\mathcal{L}h + 2\mathcal{N}[u+\hat{u}]h, 2\mathcal{N}[\hat{h}]v)_0 \\ + (2\mathcal{N}[h]\hat{h}, \mathcal{L}v + 2\mathcal{N}[u+\hat{u}]v)_0 + o(\|h\|_V, \|\hat{h}\|_V)\,. \end{aligned} \tag{4.104}$$

Again, the strong form of the second derivative is obtained through integration by parts:

$$\begin{aligned} g/2 = \mathcal{N}^T[h]\mathcal{L}\hat{h} + 2\mathcal{N}^T[h]\mathcal{N}[u+\hat{u}]\hat{h}) + \mathcal{N}^T[\hat{h}]\mathcal{L}h + 2\mathcal{N}^T[\hat{h}]\mathcal{N}[u+\hat{u}]h) \\ + \mathcal{L}^*\mathcal{N}[h]\hat{h} + 2\mathcal{N}^T[u+\hat{u}]\mathcal{N}[h]\hat{h} + o(\|h\|_V, \|\hat{h}\|_V)\,. \end{aligned} \tag{4.105}$$

The next lemma summarises some technical results that are used in the sequel [35].

Lemma 4.25. *Consider a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ in two or three dimensions ($n = 2$ or $n = 3$) and let D_i denote the derivative with respect to the i^{th} coordinate variable. For $u \in L^2(\Omega)$, $v \in H^1(\Omega)$, the product uv is in $L^{3/2}(\Omega)$ and for $u, v, w \in H^1(\Omega)$ the product uvw is in $L^{3/2}(\Omega)$, i.e. the bilinear and trilinear mappings $(u, v) \mapsto uv$, $L^2(\Omega) \times H^1(\Omega) \to L^{3/2}(\Omega)$ and $(u, v, w) \mapsto uvw$, $H^1(\Omega)^3 \to L^{3/2}(\Omega)$ are continuous with:*

$$\|uv\|_{0,3/2} \leq c\|u\|_{0,2}\|v\|_{1,2} \quad \textit{for all } u \in L^2(\Omega) \textit{ and } v \in H^1(\Omega)\,, \tag{4.106}$$

$$\|uvw\|_{0,3/2} \leq c\|u\|_{1,2}\|v\|_{1,2}\|w\|_{1,2} \quad \textit{for all } u, v, w \in H^1(\Omega)\,. \tag{4.107}$$

Proof. The first part of the above lemma is easily proven by application of the Rellich-Kondrachov embedding Theorem B.15, followed by the Hölder inequality of Propositon B.6. In fact, statements (i) and (ii) of Theorem B.15 give, with $k = 1$ and $p = 2$ for both, 2D and 3D, the continuous embeddings $H^1(\Omega) \subset L^q(\Omega)$ for $q \in [2, 6]$. Especially for $u \in L^2(\Omega)$, $v \in H^1(\Omega) \subset L^6(\Omega)$, according to (B.4) it follows that $uv \in L^r(\Omega)$ with $1/r = 1/2 + 1/6$, that is $r = 3/2$. Accordingly, due to the embedding $H^1(\Omega) \subset L^{9/2}(\Omega)$ it

follows that for $u, v, w \in H^1(\Omega)$ the product uvw lies in $L^r(\Omega)$ with $1/r = 3 \cdot 2/9$, resulting again in $r = 3/2$. Furthermore by the Hölder inequality (B.5) the inequalities:

$$\|uv\|_{0,3/2} \leq \|u\|_{0,2}\|v\|_{0,3/2} \leq c\|u\|_{0,2}\|v\|_1 \quad \forall\, u \in L^2(\Omega),\, v \in H^1(\Omega)\,,$$
$$\|uvw\|_{0,3/2} \leq \|u\|_{0,3/2}\|v\|_{0,3/2}\|w\|_{0,3/2} \leq c\|u\|_1\|v\|_1\|w\|_1 \quad \forall\, u, v, w \in H^1(\Omega)\,,$$

are obtained in which the respective second inequalities are due to the continuity of the embedding. □

In the next lemma, the properties of G that are required for the validity of the approximation result in Theorem 4.20 are established [35]:

Lemma 4.26. *Assume that mapping G is defined by* (4.97)*. For V, Y and Z given by* (4.89) *and* (4.92)*, respectively, the following are true:*

(i) For all $u \in V$, $\lambda \in D_\lambda$ holds $D_u G(\lambda, u) \in L(V, Z)$.

(ii) The second Frechet derivative $D_u^2 G(\lambda, u)$ is bounded on bounded subsets of $\Lambda \times V$.

Proof. To prove (i), consider the strong form (4.103) of $D_u G(\lambda, u)$. Via the assumption $u \in V$, the regularity $u \in H^1(\Omega)^4$ holds. Now, each of the four identities in (4.103) consists of terms of the form either $D_i uv$ or uvw, where u, v and w belong to $H^1(\Omega)$. As $D_i u$ is certainly in $L^2(\Omega)$, part (a) of Lemma 4.25 implies that $g \in L^{3/2}(\Omega)^4 = Z$. Using (4.106) and (4.107), it follows that:

$$\|D_u G(\lambda, u)h\|_Z \leq c(\lambda, u, \hat{u})\|h\|_V\,,$$

i.e., that $D_u G(\lambda, u) \in L(V, Z)$. To prove (ii), consider the strong form (4.105) of the second Fréchet derivative which discloses $D_u^2 G(\lambda, u)$ to be a bilinear form on $V \times V$. Assume that (λ, u) belongs to a bounded subset of $D_\lambda \times V$ and let $h, \hat{h} \in V$ be arbitrary. Then, similar to the above considerations, (4.105) involves only terms of the form $D_i uv$ and uvw, where u, v and w belong to $H^1(\Omega)$, such that according to Lemma 4.25 $g \in Z$. Again, estimates (4.106) and (4.107) result in:

$$\|D_u^2 G(\lambda, u)[h, \hat{h}]\|_Z \leq c(\lambda, u, \hat{u})\|h\|_V\|\hat{h}\|_V\,,$$

where c is a polynomial function of λ, $\|u\|_V$ and $\|\hat{u}\|_V$. In combination with the fact that λ and $\|u\|_V$ are bounded subsets of $\Lambda \times V$, and that $\|\hat{u}\|_V$ is fixed, it follows that $D_u^2 G(\lambda, u)$ is a continuous and thus bounded bilinear form $V \times V \to Z$. □

This completes the verification of all assumptions of Theorem 4.20. As a result, it can be concluded that error estimates (4.100) and (4.101) hold for the least-squares FE approximation as long as problem (4.91) has a regular branch of solutions with sufficient regularity.

5 Implementation and solution of linear systems

In the preceding Chapter, a weak variational least-squares formulation for the first integral form (4.12) of the steady 2D-NS equations was introduced and analysed from essentially a theoretical point of view. Existence and uniqueness results for the weak solution have been obtained in appropriate Sobolev spaces by standard elliptic PDE tools and convergence bounds established in the common H^1 and L^2 norms, both in the linear and in the non-linear case; convergence loss in the presence of irregular boundary points is investigated also.

Here, the focus is the efficient numerical implementation of these ideas, providing details of the linearisation process, least-squares weighting, matrix structure, conditioning and solution of the resulting linear systems. It is shown that the latter lend themselves well to solution by an efficient and scalable multigrid strategy; the algebraic multigrid approach in particular is well suited as a black-box solver of large and sparse linear systems resulting from discretisations based on *unstructured* grids. For this purpose a customised AMG approach for the first integral formulation is established and its convergence properties thoroughly investigated.

5.1 Method of solution

5.1.1 Structure of the fully non-linear problem

Recall that the aim is to solve the reformulation (4.13) of the 2D, steady and incompressible NS equations based on a weak variational least-squares setting of the form (4.91). Linearisation is performed *before* least squares minimization which is *not* equivalent to performing it retrospectively. Actually the latter variant would be more complicated and involve more terms; however, Payette and Reddy [192] analysed these differences and came to the conclusion that generally the more complicated version does not lead to any faster convergence and is thus not considered further. This means, via Newton-linearisation, a sequence of the following minimisation problems is solved:

$$\begin{aligned}&\text{Find } u = (u_x, u_y, \phi_x, \phi_y) \in U, \text{ such that:}\\ &\mathcal{J}(u) := \|\mathcal{W}(\mathcal{L}u + \mathcal{N}[\tilde{u}](2u - \tilde{u}) - f)\|_0^2 \leq \mathcal{J}(v) \quad \forall\, v \in V\,,\end{aligned} \tag{5.1}$$

with $\tilde{u}$ denoting the approximate solution from the previous Newton iteration step, and involving the function spaces:

$$U = \left\{ v \in [H^1(\Omega)]^4 \,\middle|\, \mathcal{B}v = g \text{ on } \partial\Omega,\ \Lambda_i v = 0 \text{ for } i = 1, \ldots, N_c \right\}, \tag{5.2}$$

$$V = \left\{ v \in [H^1(\Omega)]^4 \,\middle|\, \mathcal{B}v = 0 \text{ on } \partial\Omega,\ \Lambda_i v = 0 \text{ for } i = 1, \ldots, N_c \right\}. \tag{5.3}$$

Note, that in the implementation the constraints $\Lambda_i u = 0$ are incorporated in the solution space U which is a slight deviation from the considerations in Chapter 4, where the constraints were ensured in a least-squares sense. Furthermore, the equations comprising the differential system are multiplied by weighting factors resulting in the above presence of the weighting operator $\mathcal{W}$, which is actually a diagonal matrix with positive entries $\sqrt{w_i}$ and $w_1, \ldots, w_4 \in \mathbb{R}_{>0}$; this provides the option to give special emphasis to selected equations, for example the continuity equation with the objective of improving mass conservation. Such a weighting does not affect the principal convergence properties of the method.

The above minimisation problem, (5.1), is equivalent to the weak formulation:

$$\begin{aligned}&\text{Find } u = (u_x, u_y, \phi_x, \phi_y) \in U, \text{ such that:}\\ &(\mathcal{W}(\mathcal{L} + 2\mathcal{N}[\tilde{u}])u, \mathcal{W}(\mathcal{L} + 2\mathcal{N}[\tilde{u}])v)_0 = (\mathcal{W}(f + \mathcal{N}[\tilde{u}]\tilde{u}), \mathcal{W}(\mathcal{L} + 2\mathcal{N}[\tilde{u}])v)_0 \quad \forall\, v \in V\,,\end{aligned}$$

which can alternatively be written as:

$$\begin{aligned}&\text{Find } u = (u_x, u_y, \phi_x, \phi_y) \in U, \text{ such that:}\\ &a(u, v) + b[\tilde{u}](u, v) = \ell[\tilde{u}](v) \quad \forall\, v \in V\,,\end{aligned} \tag{5.4}$$

in terms of the symmetric bilinear forms $a(\cdot,\cdot)$, $b[\tilde{u}](\cdot,\cdot)$ and the linear form $\ell[\tilde{u}](\cdot)$:

$$a(u, v) = (\mathcal{W}\mathcal{L}u, \mathcal{W}\mathcal{L}v)_0\,, \tag{5.5}$$

$$b[\tilde{u}](u, v) = 2(\mathcal{W}\mathcal{L}u, \mathcal{W}\mathcal{N}[\tilde{u}]v)_0 + 2(\mathcal{W}\mathcal{N}[\tilde{u}]u, \mathcal{W}\mathcal{L}v)_0 + 4(\mathcal{W}\mathcal{N}[\tilde{u}]u, \mathcal{W}\mathcal{N}[\tilde{u}]v)_0\,, \tag{5.6}$$

$$\ell[\tilde{u}](v) = (\mathcal{W}(f + \mathcal{N}[\tilde{u}]\tilde{u}), \mathcal{W}(\mathcal{L} + 2\mathcal{N}[\tilde{u}])v)_0\,. \tag{5.7}$$

A conformal discrete version of the weak formulation (5.4) is simply obtained by replacing the function spaces U and V by finite dimensional subspaces $U_h \subset U$ and $V_h \subset V$[1]. For the construction of the discrete system of equations, it is convenient to neglect boundary

[1] Chapter 4 revealed that an error in the boundary approximation has no relevant effect on the solution error as long as the boundary approximation is of sufficiently high order. Under this premise, for convenience of notation, a conformal discretisation can be assumed.

conditions initially and to include them subsequently as a final procedure. In this spirit define the auxiliary discrete subspace $H^1(\Omega) \supset \tilde{V}_h = \langle v_1, \ldots, v_n \rangle$ comprising all linear combinations of n basis functions v_i, and temporarily consider the modified weak problem (5.4) with $U_h = V_h = \tilde{V}_h^4 = \langle \nu_1, \ldots \nu_{4n} \rangle$ and:

$$\nu_k = \begin{cases} (v_k, 0, 0, 0) & \text{if } k \leq n \\ (0, v_{k-n}, 0, 0) & \text{if } n < k \leq 2n \\ (0, 0, v_{k-2n}, 0) & \text{if } 2n < k \leq 3n \\ (0, 0, 0, v_{k-3n}) & \text{if } 3n < k \leq 4n\,. \end{cases} \tag{5.8}$$

Then, in terms of the auxiliary quantities:

$$a_{xx} := (a_{ij})_{i,j=1,\ldots,n} \quad \text{with } a_{ij} = (\partial_x v_i, \partial_x v_j)_0\,, \tag{5.9}$$

$$a_{xy} := (a_{ij})_{i,j=1,\ldots,n} \quad \text{with } a_{ij} = (\partial_x v_i, \partial_y v_j)_0\,, \tag{5.10}$$

$$a_{yy} := (a_{ij})_{i,j=1,\ldots,n} \quad \text{with } a_{ij} = (\partial_y v_i, \partial_y v_j)_0\,, \tag{5.11}$$

the symmetric matrix $A = (a_{kl})_{k,l=1,\ldots 4n}$, with $a_{kl} = a(\nu_k, \nu_l)$, can be constructed as a symmetric (4×4) block matrix involving the blocks $(A_{ij})_{i,j=1,\ldots,4}$ as follows:

$$A_{11} = w_1 \eta^2 a_{yy} + (w_2 \eta^2 + w_3) a_{xx}\,, \tag{5.12}$$

$$A_{12} = w_1 \eta^2 a_{xy} + (-w_2 \eta^2 + w_3) a_{xy}^T\,, \tag{5.13}$$

$$A_{13} = -w_1 \eta a_{yy} - w_2 \eta a_{xx}\,, \tag{5.14}$$

$$A_{14} = -w_1 \eta a_{xy} + w_2 \eta a_{xy}^T\,, \tag{5.15}$$

$$A_{22} = w_1 \eta^2 a_{xx} + (w_2 \eta^2 + w_3) a_{yy}\,, \tag{5.16}$$

$$A_{23} = -w_1 \eta a_{xy}^T + w_2 \eta a_{xy}\,, \tag{5.17}$$

$$A_{24} = -w_1 \eta a_{xx} - w_2 \eta a_{yy}\,, \tag{5.18}$$

$$A_{33} = (w_1 + w_4) a_{yy} + w_2 a_{xx}\,, \tag{5.19}$$

$$A_{34} = (w_1 - w_4) a_{xy} - w_2 \eta a_{xy}^T\,, \tag{5.20}$$

$$A_{44} = (w_1 + w_4) a_{xx} + w_2 \eta a_{yy}\,. \tag{5.21}$$

In a similar way, for an arbitrary function $\gamma = \gamma(x, y)$ define the operators:

$$b_{x\mathbb{1}}[\gamma] := (b_{ij})_{i,j=1,\ldots,n} \quad \text{with } b_{ij} = (\partial_x v_i, \gamma v_j)_0\,, \tag{5.22}$$

$$b_{y\mathbb{1}}[\gamma] := (b_{ij})_{i,j=1,\ldots,n} \quad \text{with } b_{ij} = (\partial_y v_i, \gamma v_j)_0\,, \tag{5.23}$$

$$b_{\mathbb{1}\mathbb{1}}[\gamma] := (b_{ij})_{i,j=1,\ldots,n} \quad \text{with } b_{ij} = (v_i, \gamma v_j)_0\,, \tag{5.24}$$

which allow the matrix $B[\tilde{u}] = (b_{kl})_{k,l=1,\dots,n}$, with $b_{kl} = b[\tilde{u}](\nu_k, \nu_l)$, to be analogously constructed as a (4×4) block matrix involving the blocks $(B_{ij})_{i,j=1,\dots,4}$ as follows:

$$\begin{aligned} B_{11} &= -\varrho\eta \left[w_1 (b_{y\mathbb{1}}[\tilde{u}_y] + b_{y\mathbb{1}}^T[\tilde{u}_y]) + w_2 (b_{x\mathbb{1}}[\tilde{u}_x] + b_{x\mathbb{1}}^T[\tilde{u}_x]) \right] \\ &\quad + \varrho^2 \left[w_1 b_{\mathbb{1}\mathbb{1}}[\tilde{u}_y^2] + w_2 b_{\mathbb{1}\mathbb{1}}[\tilde{u}_x^2] \right] , \qquad (5.25) \\ B_{12} &= \varrho\eta \left[w_2 (b_{x\mathbb{1}}[\tilde{u}_y] + b_{y\mathbb{1}}^T[\tilde{u}_x]) - w_1 (b_{y\mathbb{1}}[\tilde{u}_x] + b_{x\mathbb{1}}^T[\tilde{u}_y]) \right] \\ &\quad + \varrho^2 (w_1 - w_2) b_{\mathbb{1}\mathbb{1}}[\tilde{u}_x \tilde{u}_y] , \qquad (5.26) \\ B_{13} &= B_{24} = \varrho w_1 b_{y\mathbb{1}}^T[\tilde{u}_y] + \varrho w_2 b_{x\mathbb{1}}^T[\tilde{u}_x] , \qquad (5.27) \\ B_{14} &= -B_{23} = \varrho w_1 b_{x\mathbb{1}}^T[\tilde{u}_y] - \varrho w_2 b_{y\mathbb{1}}^T[\tilde{u}_x] , \qquad (5.28) \\ B_{22} &= -\varrho\eta \left[w_1 (b_{x\mathbb{1}}[\tilde{u}_x] + b_{x\mathbb{1}}^T[\tilde{u}_x]) + w_2 (b_{y\mathbb{1}}[\tilde{u}_y] + b_{y\mathbb{1}}^T[\tilde{u}_y]) \right] \\ &\quad + \varrho^2 (w_1 b_{\mathbb{1}\mathbb{1}}[\tilde{u}_x^2] + w_2 b_{\mathbb{1}\mathbb{1}}[\tilde{u}_2^2]) , \qquad (5.29) \\ B_{33} &= B_{34} = B_{44} = 0 . \qquad (5.30) \end{aligned}$$

The right-hand side vector in (5.4) is composed of sink and source terms from f as well as terms resulting from linearisation, i.e. $\mathcal{N}[\tilde{u}]\tilde{u}$; thus it is convenient to employ the re-definition $f \leftarrow f + \mathcal{N}[\tilde{u}]\tilde{u}$. If again auxiliary operators:

$$\begin{aligned} d_x[\gamma] &:= (d_i)_{i=1,\dots 4} \quad \text{with } d_i = (\partial_x v_i, \gamma)_0 , \qquad (5.31) \\ d_y[\gamma] &:= (d_i)_{i=1,\dots 4} \quad \text{with } d_i = (\partial_x v_i, \gamma)_0 , \qquad (5.32) \\ d_{\mathbb{1}}[\gamma] &:= (d_i)_{i=1,\dots 4} \quad \text{with } d_i = (v_i, \gamma)_0 , \qquad (5.33) \end{aligned}$$

are defined depending on an arbitrary function $\gamma = \gamma(x, y)$, the right-hand side vector $d = (d_k)_{k=1,\dots,4n}$, with $d_k = \ell(\nu_k)$, is composed of the four portions $(d_i)_{i=1,\dots,4}$:

$$\begin{aligned} d_1 &= w_1 \eta d_y[f_1] - w_2 \eta d_x[f_2] + w_3 d_x[f_3] - w_1 \varrho d_{\mathbb{1}}[\tilde{u}_y f_1] + w_2 \varrho d_{\mathbb{1}}[\tilde{u}_x f_2] , \qquad (5.34) \\ d_2 &= w_1 \eta d_x[f_1] + w_2 \eta d_y[f_2] + w_3 d_y[f_3] - w_1 \varrho d_{\mathbb{1}}[\tilde{u}_x f_1] - w_2 \varrho d_{\mathbb{1}}[\tilde{u}_y f_2] , \qquad (5.35) \\ d_3 &= w_1 d_y[f_1] - w_2 d_x[f_2] + w_4 d_y[f_4] , \qquad (5.36) \\ d_4 &= w_1 d_x[f_1] + w_2 \eta d_y[f_2] - w_4 d_x[f_4] . \qquad (5.37) \end{aligned}$$

To account for boundary conditions the linear system has to be modified after each matrix/vector construction step. Accepting that the final matrix $M = A + B$ and the right-hand side d for a linear system $M_{ij} c_j = d_i$, $i, j = 1, \dots, 4n$ have been constructed and that only Dirichlet boundary conditions are prescribed, then the coefficients c_k are known for a given index set, say $c_k = g_k$ for $k \in I_b$. They can be eliminated from the system such that the remaining coefficients are obtained by solving the set of equations:

$$M_{ij} c_j = d_i - M_{ik} c_k \quad \text{for indices } i, j = 1, \dots, 4n \setminus I_b \text{ and } k \in I_b . \qquad (5.38)$$

Elimination of boundary rows and columns is essential to retain a symmetric system matrix which will allow, as demonstrated, for the use of efficient linear solvers. Note that in some cases a re-ordering of the variables might be desirable to reduce the bandwidth of the resulting system matrix, for instance by classical methods such as the Reverse-Cuthill-McKee or the Minimum-Degree algorithm [211]; this is especially important when using direct solution methods; however, less important in the context of multigrid methods as considered in Sec. 5.4, and thus not described in detail here.

The complete procedure for solving a stationary NS problem by a least-squares FEM is summarised in Algorithm 1. In order to construct the linear system, the sub-matrices (5.12)-(5.21) have to be built once for the whole iteration procedure and the sub matrices (5.25)-(5.29) as well as the sub vectors (5.34)-(5.37) have to be updated for each iteration step based on the previous solution vector. In the special case of a weighting with $w_1 = w_2$ and $w_3 = w_4$ certain terms in the block matrices vanish, leading to simplifications. However, note that in the end the complexity of the building process reduces to the construction of the auxiliary operators (5.9)-(5.11), (5.22)-(5.24) and (5.31)-(5.33); that is, the heart of the construction process relies on the assembly of these matrices. The efficient construction of such sub-matrices is treated in Sec. 5.1.3; the corresponding FE calculations are based on standard Lagrange elements and follow classical concepts such as element-wise construction, iso-parametric transformation and efficient numerical quadrature.

Algorithm 1 Navier-Stokes solution via FEM

procedure u = NavierStokes(Ω_h, w_i, g, Re)
  Get the necessary data from a given domain discretisation Ω_h.
  Define total variable index set $I = \{1, \ldots, 4n\}$ and boundary index set I_b.
  Create Stokes matrix A according to (5.9)-(5.21).
  Initialise $u^{(0)} = 0$, $m = 0$.
  while $m < 1$ or $\|u^{(m)} - u^{(m-1)}\| > \epsilon$ **do**
    Calculate matrix $B[u^{(m)}]$ according to (5.22)-(5.30).
    Calculate right-hand side $d[u^{(m)}]$ according to (5.31)-(5.37).
    Set $M_{ij}^{(m)} := A_{ij} + B_{ij}[u^{(m)}]$.
    Set coefficients $c_k^{(m+1)}$ for $k \in I_b$ via Dirichlet conditions $\mathcal{B}u = g$.
    Account for BC via $d_i^{(m)} := d_i[u^{(m)}] - M_{ik}^{(m)} c_k^{(m+1)}\,,\ i \in I \setminus I_b\,,\ k \in I_b$.
    Solve $M_{ij}^{(m)} c_j^{(m+1)} = d_i^{(m)}\,,\ i, j \in I \setminus I_b$.
    Set $u^{(m+1)}(x, y) = \sum_{i=1}^{4n} c_i^{(m+1)} \nu_i(x, y)$.
    Set $m := m + 1$.
  end while
  return $u := u^{(m)}$.
end procedure

5.1.2 Treatment of periodic problems

Below, particular problems, i.e. Couette-like and thin film flows, involving periodicity in one direction only, the x-direction, are considered. Due to the fact that periodic boundary conditions are most conveniently handled by connecting the respective endpoints of the FE grid, it is desirable to only have to deal with periodic fields; however, that is not necessarily the case for the auxiliary potential fields ϕ_1 and ϕ_2 as is easily seen. Linear combinations of (1.41), (1.42) lead to the identity:

$$-2\sigma_{ij} = -(p+U)\delta_{ij} + \eta\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \varrho u_i u_j \ , \tag{5.39}$$

with:

$$\sigma_{11} = \frac{\partial^2 \Phi}{\partial y^2} \ , \quad \sigma_{22} = \frac{\partial^2 \Phi}{\partial x^2} \ , \quad \sigma_{12} = -\frac{\partial^2 \Phi}{\partial x \partial y} \ , \tag{5.40}$$

in which $-2\sigma_{ij}$ is identical to the stress tensor apart from the occurrence of the potential energy density U. If the physical quantities u_i and p are assumed periodic in the x-direction, then the second derivatives of Φ are periodic apart from the dependency on U. In the following consider the splitting $\Phi = \Phi_p + \Phi_a$ into a periodic plus a non-periodic (aperiodic) part, respectively.

(a) Periodic channel flow

In the simpler case of Couette flows, no external forces occur, i.e. $U = 0$. Thus all second derivatives of Φ_a can be at most constant leading to $\Phi_a(x,y) = c_1 x^2 + c_2 xy + c_3 x$ and for the representations $\phi_1 = \phi_{1,p} + \phi_{1,a}$ and $\phi_2 = \phi_{2,p} + \phi_{2,a}$ to aperiodic parts:

$$\phi_{1,a}(x,y) = c_1 x \, , \quad \phi_{2,a}(x,y) = c_2 x \, ,$$

being linear in x. In a classical channel flow configuration with upper and lower walls and periodic conditions on the left and right-hand side, the two constants c_1, c_2 are not known it advance. So when connecting the FE grid points on the left- and right-hand side this additional freedom should be taken into account. Otherwise the solution process is carried out as described in Sec. 5.1.1.

(b) A simple iteration scheme for periodic film flow

Consider gravity-driven film flow down a corrugated rigid surface inclined at an angle α to the horizontal, as shown in Fig. 5.1. Along the stationary rigid surface, velocity Dirichlet conditions are imposed while at the free surface three boundary conditions are required: a kinematic condition and two dynamic conditions in which the latter take the form of

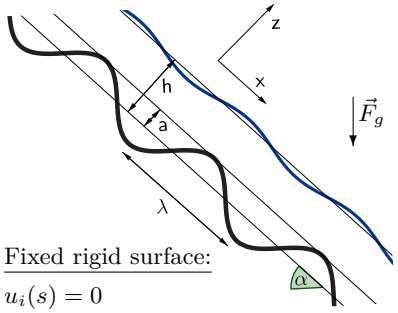


Figure 5.1: Schematic of gravity-driven film flow down a corrugated rigid surface inclined at an angle α to the horizontal. The indices i and j in the boundary conditions run from 1 to 2.

inhomogeneous Dirichlet conditions for ϕ_1 and ϕ_2 as derived in Sec. 3.2.2 of Chapter 3; these depend on the surface tension, the curvature and the potential energy density. In principle the shape of the free surface can be found by iterating over the kinematic condition while solving a sequence of flow problems with prescribed dynamic conditions in a fixed domain.

For gravity-driven flow the potential energy density takes the form $U = \varrho g[\cos(\alpha)y - \sin(\alpha)x]$ leading to a non-periodic term in (5.39). Thus the second derivatives of Φ may depend linearly on x, causing Φ to be possibly cubic in x. The non-periodic influence of U can be eliminated by subtracting a suitable known solution; for instance, the Nusselt solution for the case of a planar surface in which the free surface is correspondingly flat and of thickness h provided the Reynolds number is below a critical value [277], which is simply:

$$u_1(x,y) = \frac{\varrho g}{2\eta}\sin(\alpha)y[2h-y]\,, \quad u_2(x,y) = 0\,, \quad p(x,y) = -\varrho g\cos(\alpha)y + p_0 \tag{5.41}$$

$$\Phi(x,y) = -\frac{1}{2}\varrho g\sin(\alpha)\left[\frac{1}{2}xy^2 - hxy + \frac{1}{6}x^3\right] - \frac{p_0}{2}(x^2+y^2) + c_1x + c_2y + c_3\,, \tag{5.42}$$

in which the constants c_1, c_2, c_3 can be chosen arbitrarily; for instance the prescription $\phi_1(0,h) = \phi_2(0,h) = 0$ leads to:

$$\phi_1(x,y) = -\frac{1}{4}\varrho g\sin(\alpha)\left[x^2+y^2-2hy+h^2\right] - p_0x + c_1\,, \tag{5.43}$$

$$\phi_2(x,y) = -\frac{1}{2}\varrho g\sin(\alpha)\left[xy - hx\right] - p_0y + c_2\,, \tag{5.44}$$

with $p_0 = c_1 = c_2 = 0$. Consider now the splitting $u = (u_1, u_2, \phi_1, \phi_2) = u_p + u_a$ into a periodic plus a non-periodic part, respectively. Obviously the non-periodic part can be chosen as $u_a = u_N + u_L$ with Nusselt solution u_N and another part u_L which comprises the

remaining linear dependencies in x. The latter can be found via the dynamic boundary condition $\varepsilon_{ij}\phi_j(s) = g_i(s)$ as $u_L = (0, 0, c_1 x, c_2 x)$ with $c_i = g_i(s_1)/\lambda$ if the free surface parametrisation is such that $s \in [s_0, s_1]$. Thus, for a fixed approximation of the free surface shape a problem of the following form has to be solved:

$$\mathcal{L}u_p + 2\mathcal{N}[\tilde{u}_p + u_a]u_p = \mathcal{N}[\tilde{u}_p + u_a](\tilde{u}_p + u_a) + f\,, \tag{5.45}$$

in which the aperiodic part $f := -\mathcal{L}u_a - 2\mathcal{N}[\tilde{u}_p + u_a]u_a$ has been shifted to the right-hand side and the boundary conditions are given by:

$$u_{i,p}(s) = 0\,, \qquad \text{along the fixed rigid surface,} \tag{5.46}$$

$$\phi_{i,p}(s) = g_i(s) - \phi_{i,a}(s)\,, \qquad \text{along the free surface,} \tag{5.47}$$

for $i = 1, 2$. The above problem can be solved via the method described in Sec. 5.1.1 with the free surface updated via the kinematic boundary condition which has not been used so far. Given that an approximate solution $u^{(n)}(x, y)$ together with a surface shape, expressed in form of a height function $h^{(n)}(x)$ depending on $x = x(s)$, is available, then a correction $h^{(n+1)}(x)$ is found by shifting the surface position in direction of the (outflowing) velocity field:

$$x = \tilde{x} + \Delta t \cdot u_1^{(n)}\left(\tilde{x}, h^{(n)}(\tilde{x})\right) =: \xi^{(n)}(\tilde{x})\,, \tag{5.48}$$

$$h^{(n+1)}(x) = h^{(n)}(\tilde{x}) + \Delta t \cdot u_2^{(n)}\left(\tilde{x}, h^{(n)}(\tilde{x})\right) \qquad \text{with } \tilde{x} = \left(\xi^{(n)}\right)^{-1}(x)\,, \tag{5.49}$$

in which the points $x + \lambda z$, $z \in \mathbb{Z}$ are identified due to the λ-periodicity in the x-direction and Δt is a characteristic time scale; usually a fraction of the period time T which elapses until a surface particle has passed a period length λ. This iteration is performed until the norm difference of two consecutive height functions falls below a predefined threshold leading to the procedure shown as Algorithm 2.

A few heuristic strategies which are not mentioned in detail can be crucial to improve the stability, efficiency and convergence speed of the surface iteration. For instance the update algorithm for the surface shape is supplemented by an adaptive damping mechanism in the case of strong surface variations; first, this involves an automatic, self-adaptive choice of Δt, depending on the Reynolds and Capillary numbers (being defined later in the context of the applications), the current difference $h^{(n+1)} - h^{(n)}$ and the whole convergence history; second, a possible smoothing of a new $h^{(n+1)}$ by Fourier approximations of adaptive order. Moreover, the FE grid and the accuracy of the Newton solutions in line 6 of Algorithm 2 may be refined in the course of the surface iteration process instead of calculating high precision solutions directly from the beginning. Sometimes, instead of using (5.49), it is

more efficient to solve a differential equation of the form:

$$\frac{\mathrm{d}}{\mathrm{d}x}h^{(n+1)}(x) = \frac{{u_2}^{(n)}\left(x, h^{(n+1)}(x)\right)}{{u_1}^{(n)}\left(x, h^{(n+1)}(x)\right)}\,, \qquad h^{(n+1)}(x_0) = y_0\,, \tag{5.50}$$

in which case the streamline starting from (x_0, y_0) is traced. Since the current free surface shape and corresponding velocity field are approximations only, it may happen that a streamline $h^{(n+1)}$ starting from (x_0, h_0) lies partly outside of the current mesh; this is the reason why $h_0/2 < y_0 < h_0$ is chosen in a way that ensures the streamline remains inside the meshed domain and reaches the outflow boundary. If this is achieved, a new approximation of the free surface shape is obtained by shifting the calculated streamline $h^{(n+1)}$ by an amount $h_0 - y_0$ to the correct position. An automatic switch from (5.49) to (5.50) is implemented depending on the problem parameters and the convergence history.

Finally, it is mentioned that stationary but non-periodic film flows with standard inflow and outflow conditions can be treated in a similar way but without the need for implementing a splitting into periodic and non-periodic parts. Furthermore, in the context of a LSFEM it would be desirable to directly incorporate the free surface degrees of freedom into the system of equations and to solve the whole system in one Newton iteration; such a procedure would be similar to the "spine method" from [265] and is left as a future development.

Algorithm 2 Free surface iteration

procedure $[u,\ h]$ = FilmFlow(h, λ, σ, α, Re)
  Initialise $h^{(0)}(x) = h_0$, $m = 0$.
  while $m < 1$ or $\|h^{(n+1)}(x) - h^{(n)}(x)\| > \varepsilon$ **do**
    Calculate $g_i^{(m)}(s)$ via the dynamic BC in Fig. 5.1.
    Determine aperiodic part $u_a^{(m)} := u_N + u_L^{(m)}$.
    Solve the non-linear problem (5.45)-(5.47) with Algorithm 1.
    Set $u^{(m)} := u_p^{(m)} + u_a^{(m)}$.
    Update the free surface $h^{(m+1)}(x)$ via (5.49).
    Set $m := m + 1$.
  end while
  return $u = u^{(m-1)}$, $h = h^{(m-1)}$.
end procedure

5.1.3 Isoparametric concept

An important stage in FE calculations is the construction of the system matrix in which the course of action as well as the utilised data structures are of special interest. To set up the FE procedure for the present work the classical "p-e-t" data structure [174] is used, consisting of an array p for the node coordinates, an array e with the node indices of all

boundary edges and an array t which, for each element, indicates the node indices involved in mathematically positive order, cf. Fig. 5.3. Based on this data structure three concepts transpire as particularly useful for efficient computation: Firstly, the element-wise construction of the matrices is accepted as standard as it is much easier to handle than a nodal-based construction which requires a more complex representation of the neighbour-structure. Secondly, the element-wise integrals, possibly residing on complicated curved elements, are transformed to and calculated on reference elements, in which the transformation itself is approximated by the *same* basis functions as the solution; thus the phrase "*iso*parametric". Finally, integrals are evaluated by numerical quadrature rules which are extremely efficient and much more reasonable even if exact primitives were available; in this context Gauss-quadrature is typically used as an integration tool allowing for the best relation of exact integration order to necessary evaluation points [25].

In the following, unstructured grids comprised of both triangular and quadrilateral elements – see Fig. 5.2 – are considered in the plane and the test and solution spaces involve standard Lagrange elements up to quadratic order. The basic ideas are summarised following which the practical implementation is briefly exemplified. Hereby the concepts and the implementation are closely related to those of Bartels et al. [21].

(a) Domain decomposition and reference transformations

It is assumed that $\overline{\Omega}$ is decomposed into finitely many FE domains $T \in \mathcal{T}$ with curved boundaries which have either three or four edges, or sides, respectively. Furthermore, the existence of $\mathcal{T}_3, \mathcal{T}_4 \subseteq \mathcal{T}$ is assumed such that $\mathcal{T}_3 \cup \mathcal{T}_4 = \mathcal{T}$ and $\mathcal{T}_3 \cap \mathcal{T}_4 = \emptyset$. In order to guarantee that neighbouring elements match each other, element edges are defined through a reference parametrisation. If A and B are the endpoints of an edge E which may be curvilinear with a point C on E, then E is given by the parametrisation [21]:

$$\Phi_E : E_{\mathrm{ref}} \to \mathbb{R}^2\,, \quad t \mapsto A(1-t)/2 + B(1+t)/2 + \tilde{C}(1-t)(1+t)\,,$$

in which $\tilde{C} = C - (A+B)/2$ and $E_{\mathrm{ref}} = [-1,1]$ as in Fig. 5.2. Assume that the restriction of Φ_E to $(-1,1)$ is an immersion; this is guaranteed if A, B, and C are distinct and either C lies on the line segment connecting A and B or A, B, and C are not co-linear.

Given any triangular element $T \in \mathcal{T}_3$, the three corner vertices are mandatory and denoted by $p_1^{(T)}$, $p_2^{(T)}$, $p_3^{(T)}$, whereas the edge vertices $p_4^{(T)}$, $p_5^{(T)}$, $p_6^{(T)}$ are "optional" and can be prescribed in domain areas where a higher polynomial degree is reasonable or, for instance, on edges of curved boundaries. For the representation of a triangular element $T \in \mathcal{T}_3$ a reference element T_{ref} and form functions $N_{t1}, \ldots, N_{t6} \in H^1(T_{\mathrm{ref}})$ are defined such that each triangular element is the image of the map [21]:

$$\Phi_T = \sum_{j=1}^{6} c_j^{(T)} N_{tj} : T_{\mathrm{ref}} \to \mathbb{R}^2\,, \quad T \in \mathcal{T}_3\,. \tag{5.51}$$

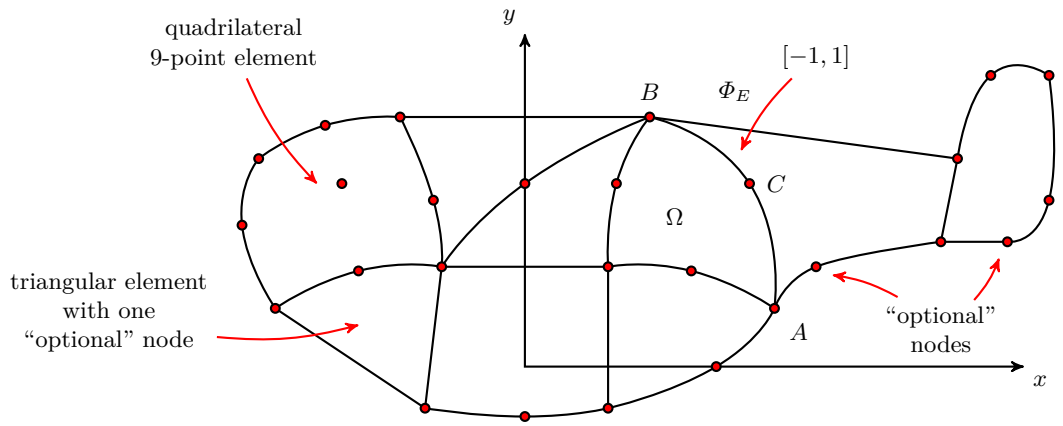


Figure 5.2: Example of an admissible decomposition of a domain involving both triangular and quadrilateral elements with different numbers of vertices. The edge from point A over C to B is mapped to the interval $[-1,1]$ via Φ_E.

Here, the coefficients $c_j^{(T)}$ are related to the given vertices of the element such that:

$$
\begin{aligned}
c_j^{(T)} &= p_j^{(T)}\,, \quad \text{for } j = 1,2,3 \\
c_4^{(T)} &= p_4^{(T)} - \frac{p_1^{(T)} + p_2^{(T)}}{2}\,, \qquad c_5^{(T)} = p_5^{(T)} - \frac{p_2^{(T)} + p_3^{(T)}}{2}\,, \\
c_6^{(T)} &= p_6^{(T)} - \frac{p_3^{(T)} + p_1^{(T)}}{2}\,,
\end{aligned}
$$

in which $c_{j+3}^{(T)} = 0$, $j = 1,2,3$ in the case that the optional edge vertices are not prescribed. The reference element is chosen as $T_{\text{ref}} := \{(\xi,\eta) \in \mathbb{R}^2 : \xi,\eta \geq 0, \xi+\eta \leq 1\}$ and the form functions defined according to:

$$
\begin{aligned}
N_{t1}(\xi,\eta) &:= 1-\xi-\eta\,, & N_{t2}(\xi,\eta) &:= \xi\,, && (5.52)\\
N_{t3}(\xi,\eta) &:= \eta\,, & N_{t4}(\xi,\eta) &:= 4\xi(1-\xi-\eta)\,, && (5.53)\\
N_{t5}(\xi,\eta) &:= 4\xi\eta\,, & N_{t6}(\xi,\eta) &:= 4\eta(1-\xi-\eta)\,. && (5.54)
\end{aligned}
$$

Given any quadrilateral element $T \in \mathcal{T}_4$, the four corner vertices are again mandatory and given as $p_j^{(T)}$, $j = 1,\ldots,4$, whereas the edge vertices $p_j^{(T)}$, $j = 5,\ldots,8$ and an additional vertex $p_9^{(T)}$ in the interior of the element are optional. For the representation of an element $T \in \mathcal{T}_4$ a reference element Q_{ref} is again defined together with form functions $N_{q1},\ldots,N_{q9} \in H^1(Q_{\text{ref}})$ such that each quadrilateral element is the image of the map [21]:

$$
\Phi_T = \sum_{j=1}^{9} d_j^{(T)} N_{qj} : Q_{\text{ref}} \to \mathbb{R}^2\,, \quad T \in \mathcal{T}_4\,. \tag{5.55}
$$

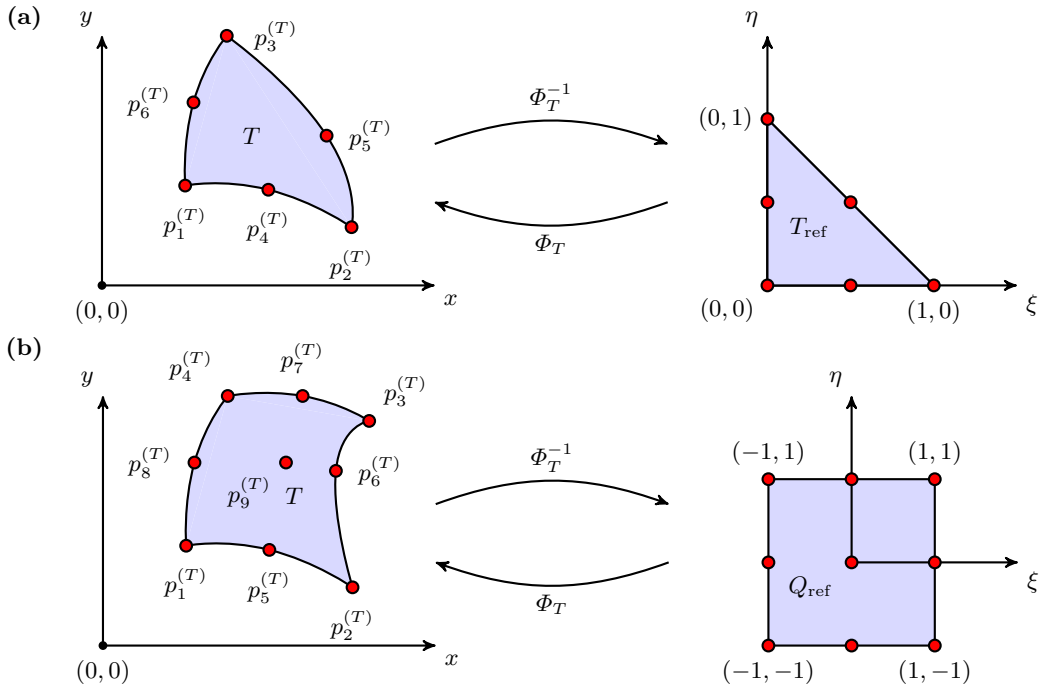


Figure 5.3: Diffeomorphisms from the reference elements T_{ref}, Q_{ref} in the (ξ, η)-plane to curved elements T in the (x, y)-domain Ω, for (a) triangular and (b) quadrilateral elements.

In this case the coefficients $d_j^{(T)}$ are related to the given vertices of the element such that:

$$
\begin{aligned}
d_j^{(T)} &= p_j^{(T)}\,, \quad \text{for } j = 1, \ldots, 4 \\
d_5^{(T)} &= p_5^{(T)} - \frac{p_1^{(T)} + p_2^{(T)}}{2}\,, \qquad d_6^{(T)} = p_6^{(T)} - \frac{p_2^{(T)} + p_3^{(T)}}{2} \\
d_7^{(T)} &= p_7^{(T)} - \frac{p_3^{(T)} + p_4^{(T)}}{2}\,, \qquad d_8^{(T)} = p_8^{(T)} - \frac{p_4^{(T)} + p_1^{(T)}}{2} \\
d_9^{(T)} &= p_9^{(T)} + \frac{1}{4}\sum_{j=1}^{4} p_j^{(T)} - \frac{1}{2}\sum_{j=5}^{8} p_j^{(T)}\,,
\end{aligned}
$$

in which $d_{j+4}^{(T)} = 0$, $j = 1, \ldots, 5$ in the case that the optional vertices are not prescribed. For $Q_{\text{ref}} := [-1, 1]^2$ the form functions are defined as follows:

$$
\begin{aligned}
N_{q1}(\xi, \eta) &:= (1 - \xi)(1 - \eta)/4\,, & N_{q2}(\xi, \eta) &:= (1 + \xi)(1 - \eta)/4\,, \\
N_{q3}(\xi, \eta) &:= (1 + \xi)(1 + \eta)/4\,, & N_{q4}(\xi, \eta) &:= (1 - \xi)(1 + \eta)/4\,, \\
N_{q5}(\xi, \eta) &:= (1 - \xi^2)(1 - \eta)/2\,, & N_{q6}(\xi, \eta) &:= (1 + \xi)(1 - \eta^2)/2\,, \\
N_{q7}(\xi, \eta) &:= (1 - \xi^2)(1 + \eta)/2\,, & N_{q8}(\xi, \eta) &:= (1 - \xi)(1 - \eta^2)/2\,, \\
N_{q9}(\xi, \eta) &:= (1 - \xi^2)(1 - \eta^2)\,.
\end{aligned}
$$

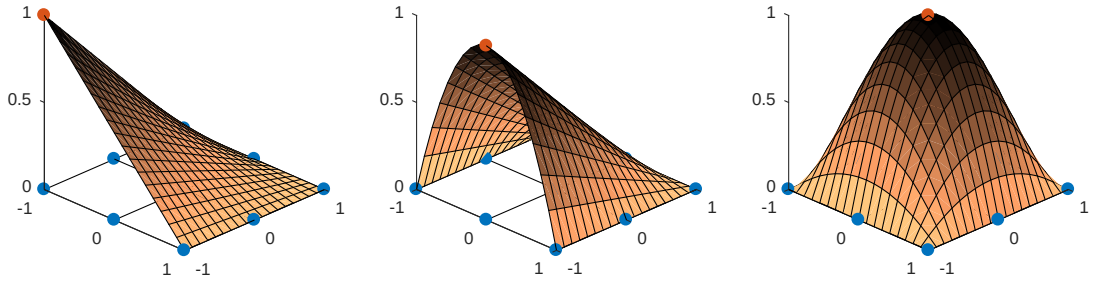


Figure 5.4: Selected shape functions on a quadrilateral reference element. From left to right: N_{q1}, N_{q5} and N_{q9}.

The reference transformations for triangular and quadrilateral elements are visualised in Fig. 5.3 and selected shape functions are shown in Fig. 5.4.

With the help of the diffeomorphism Φ_T according to (5.51) and (5.55) for $T \in \mathcal{T}_3$ and $T \in \mathcal{T}_4$, respectively, and the form functions $N_{t1}, \dots N_{t6}$ and $N_{q1}, \dots, N_{q9}$, a discrete subspace $\tilde{V}_h \subseteq H^1(\Omega)$ can be defined as follows. Let $\mathcal{P}$ denote the set of grid vertices or nodes, then for each node a basis function is defined which is not equal to zero only on the neighbouring elements containing this node. Given a node $z \in \mathcal{P}$ and an element $T \in \mathcal{T}$ with associated node indices $j \in J_T$ such that $z = p_j^{(T)}$, a piecewise definition of basis functions v_z is given by [21]:

$$v_z|_T := \begin{cases} N_{tj} \circ \Phi_T^{-1} & \text{if } z \in T \text{ and } T \in \mathcal{T}_3\,, \\ N_{qj} \circ \Phi_T^{-1} & \text{if } z \in T \text{ and } T \in \mathcal{T}_4\,, \\ 0 & \text{if } z \notin T\,, \end{cases} \tag{5.56}$$

in which $v_z \in H^1(\Omega)$ is easily checked. The solution space:

$$\tilde{V}_h = \left\{ \sum_{z \in \mathcal{N}} \alpha_z v_z : \alpha_z \in \mathbb{R} \right\}, \tag{5.57}$$

from Sec. 5.1.1 is then formed as the linear combination of all these basis functions v_z, whereas the discrete versions U_h, V_h from the spaces (5.2) and (5.3) are constructed as appropriate restrictions of $\tilde{V}_h^4$, thus incorporating the Dirichlet boundary conditions.

(b) Assumptions on the decomposition

The following assumptions concerning the decomposition $\mathcal{T}$ are made, which imply restrictions on the choice of the vertices, points on the sides of elements, and points in the interior of elements [21]. The assumptions imply that the elements with three and four

vertices define a proper decomposition of Ω in the sense that edges or sides of neighbouring elements match and that the mapping Φ_T is an diffeomorphisms in the triangular and quadrilateral case.

1. a) There exist $\mathcal{T}_3, \mathcal{T}_4 \subseteq \mathcal{T}$ such that $\mathcal{T}_3 \cup \mathcal{T}_4 = \mathcal{T}$ and $\mathcal{T}_3 \cap \mathcal{T}_4 = \emptyset$.
   b) For each $T \in \mathcal{T}_3$ there exist $\{1, \ldots, 3\} \subseteq J_T \subseteq \{1, \ldots, 6\}$ and initially prescribed points $p_j^{(T)} \in \mathbb{R}^2$, $j \in J_T$, such that T is the image of T_{ref} under Φ_T.
   c) For each $T \in \mathcal{T}_4$ there exist $\{1, \ldots, 4\} \subseteq J_T \subseteq \{1, \ldots, 9\}$ and initially prescribed points $p_j^{(T)} \in \mathbb{R}^2$, $j \in J_T$, such that T is the image of Q_{ref} under Φ_T.
2. The closure of Ω is covered exactly by $\mathcal{T}$, i.e., $\overline{\Omega} = \bigcup_{T \in \mathcal{T}} T$ and the interior of the elements is non-intersecting, i.e. $\text{int}(T) \cap \text{int}(T') = \emptyset$ for all $T, T' \in \mathcal{T}$.
3. If $T \cap T' = \{x\}$ for $T, T' \in \mathcal{T}$ and some $x \in \mathbb{R}^2$ then x is a vertex of both elements T and T'.
4. If $T \cap T' = \{x, y\}$ for $T, T' \in \mathcal{T}$ and distinct $x, y \in \mathbb{R}^2$ then T and T' share an entire side.
5. There exists $c > 0$ such that $|\det D\Phi_T| > c$ for all $T \in \mathcal{T}$, in which $D\Phi_T$ denotes the Jacobian, see the next paragraph (c).

(c) Element-wise construction of the discrete system

In order to construct the system matrices A and B, the sub matrices (5.12)-(5.21) have to be built once for the whole iteration procedure and the sub matrices (5.25)-(5.29) have to be updated for each iteration step based on the previous solution vector. These sub matrices can again be compounded from the blocks (5.9)-(5.11) and (5.22)-(5.24); that is, the heart of the construction process relies on the assembly of the latter six block matrices.

Consider now the assembly of a_{xy} from (5.10) as an example keeping in mind that the remaining block matrices from (5.9)-(5.11) can be assembled in the same manner. The presentation below is again close to Bartels et al. [21]. In the case of a triangulated two-dimensional domain $\overline{\Omega} \subset \mathbb{R}^2$ involving n grid points, the discrete approximation space $\tilde{V}_h = \langle v_1, \ldots v_n \rangle$ consist of n continuous and piecewise smooth basis functions as defined in the previous section (b). The basis function $v_i(x, y)$ has a function value of 1 at the grid point with index i and a function value of 0 at any other grid point; Fig. 5.5 shows a basis function v_i for the piecewise linear case. The entries of a_{xy} are accumulated by:

$$(a_{xy})_{ij} = (\partial_x v_i, \partial_y v_j)_0 = \int_\Omega \partial_x v_i \partial_y v_j \, \mathrm{d}\Omega = \sum_{T \in \mathcal{T}} \int_T \partial_x v_i \partial_y v_j \, \mathrm{d}T \,, \tag{5.58}$$

in which $\mathcal{T}$ is the set of all elements. As $\partial_x v_i \partial_y v_j$ in the last expression is non-zero, i.e. $\text{supp}(v_i) \cap \text{supp}(v_j) \neq \emptyset$, only if the indices i and j both belong to the element T, the matrix can conveniently be assembled *element-wise*. It is sufficient to compute for each element $T \in \mathcal{T}_3$ a matrix $M^{(T)} = (M^{(T)}_{jk})_{j,k \in J_T}$ defined by:

$$M^{(T)}_{jk} = \int_T \partial_x (N_{tj} \circ \Phi_T^{-1}) \cdot \partial_y (N_{tk} \circ \Phi_T^{-1}) \, \mathrm{d}T \,,$$

and for each $T \in \mathcal{T}_4$ a matrix $M^{(T)} = (M^{(T)}_{jk})_{j,k \in J_T}$ defined by:

$$M^{(T)}_{jk} = \int_T \partial_x (N_{qj} \circ \Psi_T^{-1}) \cdot \partial_y (N_{qk} \circ \Psi_T^{-1}) \, \mathrm{d}T \,.$$

Thus, matrices $M^{(T)} \in \mathbb{R}^{6\times 6}$ and $M^{(T)} \in \mathbb{R}^{9\times 9}$ are computed for $T \in \mathcal{T}_3$ and $T \in \mathcal{T}_4$ involving the indices $j, k \in J_T$ of the corresponding elements. The element-wise construction of the system matrix by contributions from *local element matrices* $M^{(T)}$ is illustrated in Fig. 5.6. This process is easily extended to more complex and three-dimensional elements.

For a further, more compact explanation, the annotations "t" and "q" in the shape functions N_{tj}, N_{qj} are skipped and automatically assumed in the context of either a triangular or a quadrilateral element. Moreover, in order to avoid repetition the notation "T_{ref}" covers both triangular and quadrilateral reference elements. The Jacobian of the coordinate transformation Φ_T at a point $p = (\xi_m, \eta_m)$, is defined by:

$$D\Phi_T(p) := \frac{\partial(x,y)}{\partial(\xi,\eta)}(p) = \begin{bmatrix} \sum_j c^{(T)}_{j,1} \dfrac{\partial N_j}{\partial \xi}(p) & \sum_j c^{(T)}_{j,2} \dfrac{\partial N_j}{\partial \xi}(p) \\ \sum_j c^{(T)}_{j,1} \dfrac{\partial N_j}{\partial \eta}(p) & \sum_j c^{(T)}_{j,2} \dfrac{\partial N_j}{\partial \eta}(p) \end{bmatrix} , \tag{5.59}$$

and its inverse denoted by $D^{-1}\Phi_T$, in which single matrix entries of the inverse Jacobian

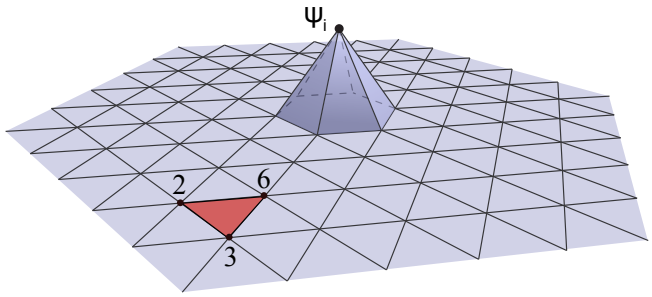


Figure 5.5: Linear basis function ψ_i and selected triangular element with corner point indices 2, 3 and 6.

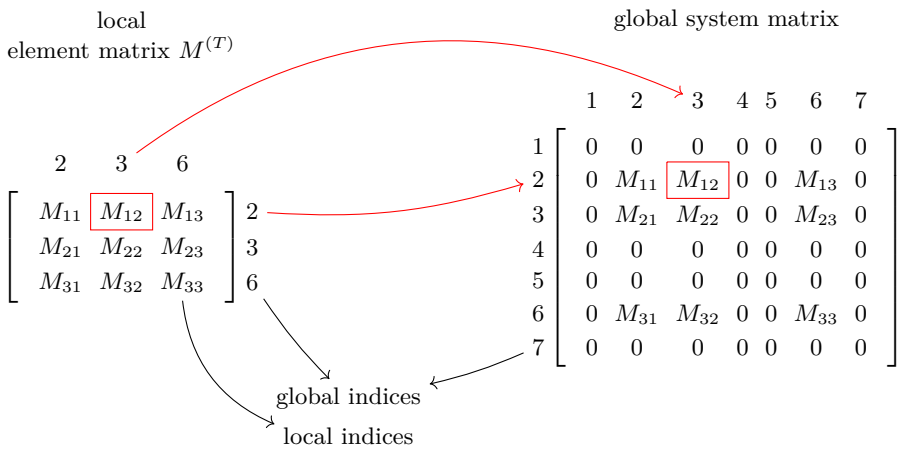


Figure 5.6: Assembly of the global system matrix out of local element matrices.

are denoted by $D_{ij}^{-1}\Phi_T$. Employing the substitution rule for the diffeomorphism Φ_T and using the identity $(D\phi_T^{-1}) \circ \Phi_T = (D\Phi_T)^{-1}$ yields:

$$\begin{aligned} M_{jk}^{(T)} &= \int_T \partial_1(N_j \circ \Phi_T^{-1}) \cdot \partial_2(N_k \circ \Phi_T^{-1}) \,\mathrm{d(x,y)} \\ &= \int_{T_{\mathrm{ref}}} \left[D_{1l}^{-1}\Phi_T(\xi,\eta)\partial_l N_j(\xi,\eta)\right] \left[D_{2l}^{-1}\Phi_T(\xi,\eta)\partial_l N_k(\xi,\eta)\right] \left|\det D\Phi_T(\xi,\eta)\right| \mathrm{d}(\xi,\eta) \end{aligned}$$

in which the Einstein summation convention has been used. In order to evaluate $D\Phi_T$, temporarily missing, i.e. not initially specified, points $p_{j+4}^{(T)}$, $j = 1,\ldots,5$ are computed by interpolation. The local stiffness matrix is then approximated using a quadrature rule on T_{ref} which is defined by points $(\xi_m,\eta_m) \in T_{\mathrm{ref}}$ and weights γ_m for $m = 1,\ldots,K_T$:

$$\begin{aligned} M_{jk}^{(T)} \approx \sum_{m=1}^{K_T} \gamma_m &\left[D_{1l}^{-1}\Phi_T(\xi_m,\eta_m)\partial_l N_j(\xi_m,\eta_m)\right] \\ &\cdot \left[D_{2l}^{-1}\Phi_T(\xi_m,\eta_m)\partial_l N_k(\xi_m,\eta_m)\right] \left|\det D\Phi_T(\xi_m,\eta_m)\right| . \end{aligned} \tag{5.60}$$

In the implementation Gaussian quadrature is employed in which quadrature points and weights for reference triangle and quadrilateral elements are taken from Bathe [25]; see also Fig. 5.7.

In the same way that a_{xy}, from (5.58), is assembled by element matrices according to (5.60), the block matrices a_{xx} and a_{yy}, from (5.9) and (5.11), can be assembled; this only needs modification of the corresponding indices. Slightly differently but in a completely analogous way, the matrices (5.22)-(5.24) can be assembled which is exemplified by $b_{x\mathbb{1}}[f]$ for an arbitrary function $f = f(x,y)$. The corresponding calculations:

$$(b_{x\mathbb{1}})_{ij} = (\partial_x v_i, f v_j)_0 = \int_\Omega \partial_x v_i f v_j \,\mathrm{d}\Omega = \sum_{T\in\mathcal{T}} \int_T \partial_x v_i f v_j \,\mathrm{d}T\,, \tag{5.61}$$

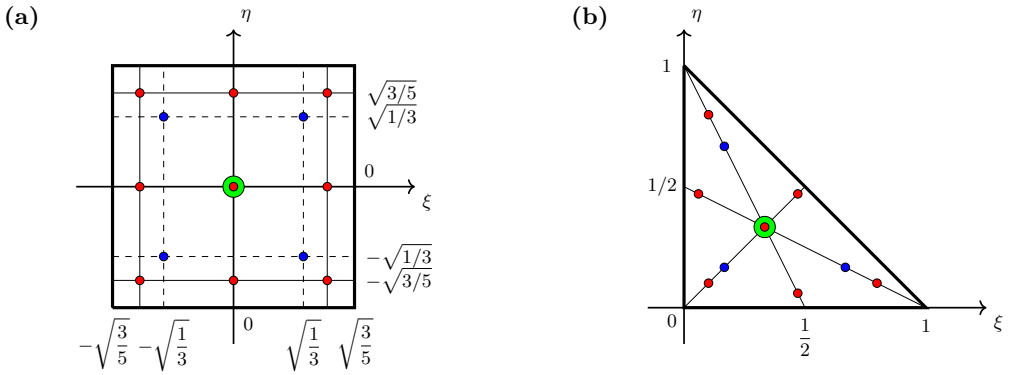


Figure 5.7: (a) Gaussian quadrature on the unit square involving one (green), four (blue) and nine (red) quadrature points. (b) Gaussian quadrature on the reference triangle involving one (green), three (blue) and seven (red) quadrature points.

require the allocation of element matrices:

$$\begin{aligned} M_{jk}^{(T)} &= \int_T \partial_1(N_j \circ \Phi_T^{-1})(x,y) \cdot f(x,y) N_k \circ \Phi_T^{-1}(x,y)\,\mathrm{d(x,y)} \\ &= \int_{T_{\mathrm{ref}}} \left[D_{1l}^{-1}\Phi_T(\xi,\eta)\partial_l N_j(\xi,\eta)\right] \left[f(\Phi_T(\xi,\eta))N_k(\xi,\eta)\right] \left|\det D\Phi_T(\xi,\eta)\right| \mathrm{d}(\xi,\eta) \\ &\approx \sum_{m=1}^{K_T} \gamma_m \left[D_{1l}^{-1}\Phi_T(\xi_m,\eta_m)\partial_l N_j(\xi_m,\eta_m)\right] \\ &\qquad \cdot \left[f(\Phi_T(\xi_m,\eta_m))N_k(\xi_m,\eta_m)\right] \left|\det D\Phi_T(\xi_m,\eta_m)\right|. \end{aligned} \tag{5.62}$$

Finally, vectors of the form (5.31)-(5.33), including parts of the linearisation as well as volume forces, are required. Their calculation is exemplified via $d_x[f]$, again for an arbitrary function $f(x,y)$. The corresponding calculations:

$$(d_x)_i = (\partial_x v_i, f)_0 = \int_\Omega f \partial_x v_i \,\mathrm{d}\Omega = \sum_{T\in\mathcal{T}} \int_T f \partial_x v_i \,\mathrm{d}T\,, \tag{5.63}$$

require the allocation of element vectors:

$$\begin{aligned} M_j^{(T)} &= \int_T \partial_1(N_j \circ \Phi_T^{-1})(x,y) \cdot f(x,y)\,\mathrm{d(x,y)} \\ &= \int_{T_{\mathrm{ref}}} \left[D_{1l}^{-1}\Phi_T(\xi,\eta)\partial_l N_j(\xi,\eta)\right] \left[f(\Phi_T(\xi,\eta))\right] \left|\det D\Phi_T(\xi,\eta)\right| \mathrm{d}(\xi,\eta) \\ &\approx \sum_{m=1}^{K_T} \gamma_m \left[D_{1l}^{-1}\Phi_T(\xi_m,\eta_m)\partial_l N_j(\xi_m,\eta_m)\right] \cdot \left[f(\Phi_T(\xi_m,\eta_m))\right] \left|\det D\Phi_T(\xi_m,\eta_m)\right|. \end{aligned}$$

(d) Boundary estimates for isoparametric finite elements

At this point attention is returned to Assumptions 4.6 concerning the boundary approximation used in the FE method. Lemma 5.2 below demonstrates why these assumptions are fulfilled for the isoparametric elements under consideration. Prior to this, some well-known results concerning inverse Sobolev inequalities, taken from Brenner and Scott [48], are cited:

Lemma 5.1. *Let $\rho h \leq \operatorname{diam}(T) \leq h$, be the diameter of an element T with incircle radius ρh where $0 < h \leq 1$, and $\mathcal{P}$ be a finite-dimensional subspace of $W_p^l(T) \cap W_q^m(T)$, where $1 \leq p, q \leq \infty$ and $0 \leq m \leq l$. Then there exists $c = c(\hat{\mathcal{P}}, \hat{\mathcal{T}}, l, p, q, \rho)$ such that for all $v \in \mathcal{P}$, we have:*

$$\|v\|_{W_p^l(T)} \leq c h^{m-l+n/p-n/q} \|v\|_{W_q^m(T)} \,. \tag{5.64}$$

Here, W_p^l denotes the classical Sobolev spaces, cf. Definition B.9 in Appendix B, where $H^l = W_2^l$. The above Lemma allows for the estimation of higher order Sobolev norms against lower order norms with an appropriate h-dependent factor; a result that is needed in the proof of the following estimates:

Lemma 5.2. *Consider a Lagrangian isoparametric finite element space $\tilde{V}_h \subset H^1(\mathcal{T})$ of the form* (5.56), (5.57) *but with arbitrary polynomial order k. Let $\mathcal{T}$ denote the corresponding triangulation with elements $T_i \in \mathcal{T}$ which approximates a given domain $\Omega \subset \mathbb{R}^2$ such that the boundary approximation $\partial\mathcal{T}$ of $\partial\Omega$ consists of transformed piece-wise k-th order polynomial curves. Furthermore, define by $u \in V_h$ elements of the product space:*

$$V_h := \{u \in \tilde{V}_h^4 \,|\, \mathcal{B}(x)u(x) = 0 \text{ for all } x \in \partial\mathcal{T}\} \,,$$

with one of the boundary operators given in (4.24), (4.25)[2]. *Then there exist constants $h_1 > 0$ and $c > 0$ such that for all $u \in V_h$, $0 < h \leq h_1$ holds:*

$$\|\mathcal{B}u\|_{0,\partial\Omega} \leq c h^{(2k+1)/2} \|u\|_{1,\Omega} \,, \tag{5.65}$$

$$\|\mathcal{B}u\|_{1,\partial\Omega} \leq c h^{(2k-1)/2} \|u\|_{1,\Omega} \,. \tag{5.66}$$

If additionally $u \in H^2(\Omega)^4$, then the norms $\|u\|_{1,\Omega}$ in (5.65) *and* (5.66) *can alternatively be replaced by $h^{1/2}\|u\|_{2,\Omega}$.*

Proof. Boundary integrals are estimated using a typical method for non-conforming FEs [48] with the analysis confined to triangular elements, although a generalisation to quadrilateral elements is straight-forward. The essential analysis is performed in the reference

[2] Note, that by extension with the zero solution and restriction of $u \in V_h$ to Ω, then without loss of generality it also holds that $u \in H^1(\Omega)^4$.

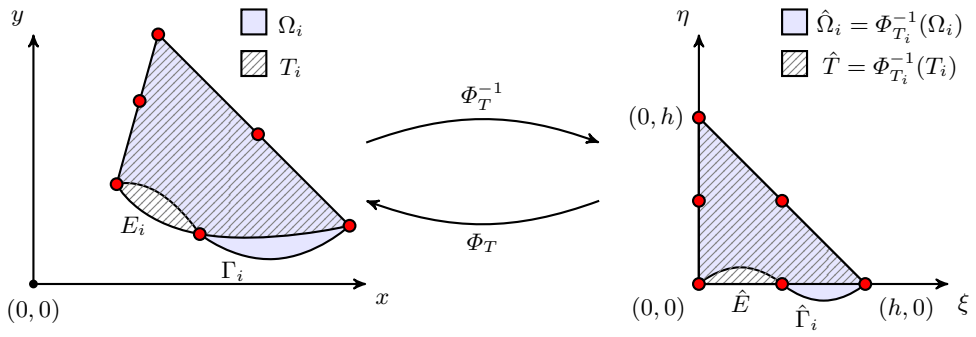


Figure 5.8: Schematic of the transformation of an element (left) located at the boundary to a reference domain (right) and vice versa. The curved element edge E_i is transformed to the straight line $\hat{E}$ while the corresponding part of the exact boundary Γ_i is generally transformed to the curve $\hat{\Gamma}_i = \Phi_{T_i}^{-1}(\Gamma_i)$.

domain and then transferred back to the computational domain; this requires the precise assignment of different edge and domain quantities. Consider a boundary element $T_i \in \mathcal{T}$ with its adjacent boundary edge $E_i \subset \partial\mathcal{T}$, then the corresponding part of the exact domain is denoted by $\Omega_i \subset \Omega$ with boundary edge $\Gamma_i \subset \partial\Omega$, cf. Fig. 5.8. An isoparametric mapping $\Phi_{T_i}^{-1} : T_i \to \hat{T}$ is used which transforms element T_i to reference element $\hat{T}$ having straight edges and corner points $(0,0)$, $(h,0)$ and $(0,h)$ in the reference coordinate system (ξ, η); the transformed quantities are denoted with a hat:

$$\Phi_{T_i}(\hat{T}) = T_i\,, \qquad \Phi_{T_i}(\hat{\Omega}_i) = \Omega_i\,,$$
$$\Phi_{T_i}(\hat{E}) = E_i\,, \qquad \Phi_{T_i}(\hat{\Gamma}_i) = \Gamma_i\,.$$

in which $\hat{E}$ can assumed to be the straight edge from the coordinate system origin to the point $(\xi, \eta) = (h, 0)$, without loss of generality. Furthermore $\hat{\Gamma}_i$ can be parametrized as $(\xi, \eta(\xi))$ for $0 \le \xi \le h$ with $\eta(\xi_j) = 0$ for $k+1$ fixed points $\xi_i \in [0, h]$. A parametric approximation of degree k gives:

$$|\eta(\xi)| \le c h^{k+1} \quad \text{and} \quad |\eta'(\xi)| \le c h^k\,, \tag{5.67}$$

for all $\xi \in [0, h]$ due to the fact that $\eta(\xi)$ can be interpreted as an interpolation error for a polynomial of degree k and, similarly, $\eta'(\xi)$ for a polynomial of degree $k-1$, cf. [31].

Consider first the case of a homogeneous Dirichlet boundary condition, (4.24), with, for instance, $\mathcal{B}u(x) = (u_1(x), u_2(x)) = (0,0)$ for all $x \in \partial\mathcal{T}$. Here, it is sufficient to examine one of the components $v := u_i$ and its transform $\hat{v} = \Phi_{T_i} v$. A Taylor series expansion leads to:

$$\hat{v}(\xi, \eta(\xi)) = \sum_{l=1}^{k} \frac{\partial_\eta^{(l)} \hat{v}(\xi, 0)}{l!} \eta^l(\xi)\,, \tag{5.68}$$

in which the index $l = 0$ is skipped due to $v(\xi, 0) = 0$. Using expressions (5.67), (5.68) and the inverse estimates (5.64) the following estimate is derived for an arbitrary transformed boundary element T_i and a function $\hat{v}$:

$$\begin{aligned}
\|\hat{v}\|^2_{0,\hat{\Gamma}_i} &= \int_{\hat{\Gamma}_i} |\hat{v}(s)|^2 \,\mathrm{d}s = \int_0^h \hat{v}^2(\xi, \eta(\xi)) \sqrt{1 + \eta'(\xi)^2} \,\mathrm{d}\xi \\
&\leq 2 \int_0^h \hat{v}^2(\xi, \eta(\xi)) \,\mathrm{d}\xi = 2 \int_0^h \left(\sum_{l=1}^{k} \frac{\partial_\eta^{(l)} \hat{v}(\xi, 0)}{l!} \eta^l(\xi) \right)^2 \mathrm{d}\xi \\
&\lesssim \sum_{l=1}^{k} h^{2l(k+1)} \int_0^h \left[\partial_\eta^{(l)} \hat{v}(\xi, 0) \right]^2 \mathrm{d}\xi \leq \sum_{l=1}^{k} h^{2l(k+1)} \|\hat{v}\|^2_{l,\hat{E}} \\
&\lesssim \sum_{l=1}^{k} h^{2(lk+1)} \|\hat{v}\|^2_{1,\hat{E}} \leq k h^{2(k+1)} \|\hat{v}\|^2_{1,\hat{E}} \lesssim h^{2k+1} \|\hat{v}\|^2_{1,\hat{T}} \,. \qquad (5.69)
\end{aligned}$$

In arriving at the last line, (5.69), Lemma 5.1 has been used with $p = q = 2$ and $m = 1$, $l = 1, \ldots, k$. Moreover, in the very last inequality use is made of the norm equivalence relations:

$$\|v\|_{0,\hat{T}} \eqsim h \|v\|_{0,\hat{E}} \,, \qquad (5.70)$$

which hold for all $v \in V_h$ due to scaling arguments and the fact that the considered space, restricted to an element $\hat{T}$, is of finite dimension [48, 160]. Following the same argument:

$$\|v\|_{0,\hat{T}} \eqsim \|v\|_{0,\hat{\Omega}_i} \,, \qquad (5.71)$$

holds for all $v \in V_h$ if $\hat{T}$ is sufficiently close to $\hat{\Omega}_i$ such that it allows invertible mapping between $\hat{T}$ and $\hat{\Omega}_i$.

The above result can be transferred from the reference to the computational domain via the following inequalities with $m = r = 1$, cf. [29]:

$$\|v\|_{0,\Gamma_i} \leq c \|\hat{v}\|_{0,\hat{\Gamma}_i} \,, \qquad (5.72)$$

$$\|\hat{v}\|_{m,\hat{\Omega}_i} \leq \| \det J_{\Phi_{T_i}} \|^{-1/2}_{L^\infty(\hat{\Omega}_i)} \|J_{\Phi_{T_i}}\|^m_{L^\infty(\hat{\Omega}_i)} \sum_{r=0}^{m} \|v\|_{r,\Omega_i} \,. \qquad (5.73)$$

Obviously $\Omega = \bigcup \Omega_i$ holds, but note that a summation over all boundary edges gives $\|v\|_{0,\partial\Omega} \leq c h^{k+1/2} \|v\|_{1,\Omega_\Gamma}$ involving the sub domain $\Omega_\Gamma \subset \Omega$ which includes only those Ω_i that meet the boundary at two points at least. Clearly, this directly implies the desired result (5.65) for Dirichlet boundary conditions; if, however, the additional condition $v \in H^2(\Omega)$ is assumed, then the Poincaré-Friedrich inequality (B.10):

$$\|v\|_{1,\Omega_\Gamma} \leq s |v|_{2,\Omega_\Gamma} \leq s \|v\|_{2,\Omega} \,,$$

holds with $s = |\Omega_\Gamma|^{1/2} = ch^{1/2}$ such that the term $\|u\|_{1,\Omega}$ in (5.65) and (5.66) can alternatively be replaced by $h^{1/2}\|u\|_{2,\Omega}$.

Using the same arguments as above the following estimate is obtained:

$$\begin{aligned}\int_{\hat\Gamma_i} |\hat v'(s)|^2 \,\mathrm{d}s &= \int_0^h \frac{|\hat v'(\xi,\eta(\xi))|^2}{\sqrt{1+\eta'(\xi)^2}}\,\mathrm{d}\xi \le \int_0^h |\hat v'(\xi,\eta(\xi))|^2\,\mathrm{d}\xi \\ &\lesssim \sum_{l=1}^{k-1} h^{2l(k+1)} \int_0^h \left[\partial_\eta^{(l)}\partial_\xi \hat v(\xi,0)\right]^2 \,\mathrm{d}\xi \\ &\le \sum_{l=1}^{k-1} h^{2l(k+1)} \|\hat v\|^2_{l+1,\hat E} \lesssim \sum_{l=1}^{k-1} h^{2lk}\|\hat v\|^2_{1,\hat E} \\ &\le (k-1)h^{2k}\|\hat v\|^2_{1,\hat E} \lesssim h^{2k-1}\|\hat v\|^2_{1,\hat T} \lesssim h^{2k-1}\|\hat v\|^2_{1,\hat\Omega_i}\,,\end{aligned}$$

and therefore:

$$\|\hat v\|^2_{1,\hat\Gamma_i} \lesssim \int_{\hat\Gamma_i} |\hat v(s)|^2 + |\hat v'(s)|^2\,\mathrm{d}s \lesssim h^{2k-1}\|\hat v\|^2_{1,\hat T}\,.$$

As before, the transformation estimate (5.73) and the equivalence relations (5.71) together with a summation over all boundary edges gives $\|v\|_{1,\partial\Omega} \le ch^{k-1/2}\|v\|_{1,\Omega_\Gamma}$. Thus the result (5.66) for Dirichlet boundary conditions follows immediately and the additional estimate for $\|\cdot\|_{2,\Omega}$ follows by the same arguments as above. The corresponding results for the remaining boundary operators (4.25) can be derived analogously; the proof of which is omitted.

□

5.2 Conditioning of linear systems

It is desirable to determine bounds for the condition number of the system matrices that arise, which is crucial for the assessment of possible linear solvers. For this purpose considerations by Braess [44] are extended to the present case of Petrovskii systems. Consider solutions $u \in U_h$ of problem (5.4) with U, V replaced by the discrete Lagrangian finite element spaces U_h, V_h as described in Sec. 5.1.3(a). For convenience a polygonal domain $\Omega \subset \mathbb{R}^2$ is assumed such that the boundary approximation is exact and only non-curved, straight elements are involved in the triangulation; in this case the FE discretisation is *conforming*. For a domain discretisation with N given grid points the functions of the nodal basis v_i are normed such that:

$$v_i(z_j) = h^{-1}\delta_{ij}\,,$$

for $i, j = 1, \ldots, N$ involving the Kronecker delta. In order to describe the conditioning of the resulting linear systems it is necessary to introduce discrete vector norms in contrast to the above Sobolev norms. In the following, the Euclidean vector norm is denoted by $\|x\|_e^2 = (x, x)_e := x^T x$ for any $x \in \mathbb{R}^n$. Whenever $\|u\|_e$ or $(u, u)_e$ is written for a function $u = \sum_{i=1}^{4N} c_i \nu_i \in H^1(\Omega)^4$, then, u is identified with the nodal vector $(c_1, \ldots, c_{4N})$ corresponding to the basis $\{\nu_i\}_{i=1,\ldots 4N}$ from (5.8), that is:

$$\|u\|_e^2 := \sum_{i=1}^{4} \sum_{j=1}^{N} u_i(z_j)^2 = \sum_{i=1}^{4N} c_i^2 \,.$$

This is reasonable since, with a h-independent constant c, the following holds:

$$c^{-1}\|u\|_e \leq \|u\|_{0,\Omega} \leq c\|u\|_e \,, \tag{5.74}$$

and therefore the Euclidean norm is equivalent to the $L^2(\Omega)$ norm.

In order to encapsulate the main idea in a brief and convenient way, further consideration is subsequently restricted to the linear Stokes flow case, although they similarly carry over to the non-linear one. Under the premise that FE quadrature is performed without error, for a FE approximation $u \in U_h$ the identity:

$$a(u, u) = (u, Au)_e \,, \tag{5.75}$$

is valid in which $a(\cdot, \cdot)$ denotes the bilinear form from (4.49), which in the proof of Corollary 4.10 has been shown to be elliptic, and A denotes the discrete system matrix of the resulting linear system. Employing (5.74), (5.75) along with the inverse inequality [44]:

$$\|u\|_1 \leq ch^{-1}\|u\|_0 \,, \tag{5.76}$$

and the continuity of the bilinear form $a(\cdot, \cdot)$ yields:

$$\lambda_{max} = \sup_{u \in U_h} \frac{(u, Au)_e}{(u, u)_e} \lesssim \sup_{u \in U_h} \frac{a(u, u)}{\|u\|_0^2} \leq \sup_{u \in U_h} \frac{\beta\|u\|_1^2}{\|u\|_0^2} \lesssim \sup_{u \in U_h} \frac{h^{-2}\|u\|_0^2}{\|u\|_0^2} = h^{-2} \,, \tag{5.77}$$

with a continuity constant β according to Definition 4.8. The simple relation $\|u\|_1 \geq \|u\|_0$ and the coerciveness of the bilinear form leads to:

$$\lambda_{min} = \inf_{u \in U_h} \frac{(u, Au)_e}{(u, u)_e} \gtrsim \inf_{u \in U_h} \frac{a(u, u)}{\|u\|_0^2} \geq \inf_{u \in U_h} \frac{\alpha\|u\|_1^2}{\|u\|_0^2} \gtrsim \inf_{u \in U_h} \frac{\|u\|_0^2}{\|u\|_0^2} = 1 \,,$$

with a coerciveness constant α. Obviously the bounds for the eigenvalues identifiy the condition of matrix A to be of magnitude $\lambda_{max}/\lambda_{min}$, that is $\mathrm{cond}(A) = \mathcal{O}(h^{-2})$. This estimate is sharp and the term h^{-2} in (5.77) is optimal which relies on the fact that the inverse inequality (5.76) is sharp. Braess [44] shows that functions $u \in U_h$ exist for which (5.76) holds with an equality symbol.

Certainly the condition number of a linear system is a relevant indicator of whether or not the system is difficult to solve. Not only can the error propagation be characterised by the condition number but also the convergence rate of an iterative solver, such as conjugate gradient (CG), the use of which is essential for large and sparse systems, Saad [211]. A growth rate of h^{-2} for the condition number for increasingly refined grids is typical for elliptic systems of Petrovskii type in general and is also typical for standard FE discretisations of the Stokes and Navier-Stokes equations; this is known for the Taylor-Hood element in the primitive variable case or for equal order Lagrange elements in the case of streamfunction-vorticity formulations [107]; so in this respect the method developed here is at least competitive.

However, as iterative solvers are almost always complemented by a suitable form of preconditioning, also the effort necessary to construct a reasonably exact inverse, for instance by multigrid techniques, plays an important role. In this context it is easily shown, along the lines of Bochev and Gunzburger [37], that the system matrix A is *spectrally equivalent* to the block diagonal matrix $\hat{D} = \mathrm{diag}(D, D, D, D)$ with:

$$D_{ij} = (v_i, v_j)_1 \, .$$

Using (5.75) and the following similar relationship:

$$(u_i, Du_i)_e = (u_i, u_i)_1 \, ,$$

reveals, together with the continuity and coerciveness of $a(\cdot, \cdot)$, that:

$$c^{-1} \sum_{i=1}^{4} (u_i, Du_i)_e \leq (u, Au)_e \leq c \sum_{i=1}^{4} (u_i, Du_i)_e \, ,$$

which means that the condition number of $\hat{D}^{-1}A$ is bounded from above by a constant independent of the matrix size and therefore independent of h.

As an important consequence of these considerations, least-squares algebraic problems for Petrovskii systems, including the one under consideration, can be preconditioned using any good preconditioner for the Poisson equation; in this respect all the geometric and algebraic multigrid methods, originally designed for Poisson-like equations, are directly accessible.

5.3 Numerical tests and validation

In order to validate the above FE method expressed in terms of the velocities and two potential fields a selection of test cases is considered which exhibit increasing complexity

with regard to the geometry and boundary conditions: this involves singly and multiply-connected domains with curved boundaries and sharp corners. Various combinations of boundary conditions are explored; pure velocity and pure potential conditions as well as different types of mixed conditions.

At first investigations are confined to Stokes flow, that is non-linear inertial effects are neglected, the aim being to verify the predicted convergence rates reported in Sec. 4.3.3(a) and (b). Full NS flow is then considered with special emphasis placed on the use of Newton iteration which is expected to converge faster when compared to corresponding primitive variable formulations, due to a reduction in the order of the non-linearity. Further aspects include the influence of grid structure on the error and the analysis of mass conservation: as the present least-squares approach assures mass conservation only globally and not element-wise, convergence is achievable but only slowly on coarse grids; resulting in non-physical results with streamlines ending at walls unless a minimum number of elements is utilised. This problem is traditionally overcome by increasing the polynomial degree or alternatively by weighting the least-squares functional and giving solution of the continuity equation priority.

5.3.1 Linear test cases and results

(a) Laminar Taylor-Couette [TF] flow

The flow between two infinitely long concentric cylinders with radii $R_1 < R_2$ and rotating at different angular velocities ω_1, ω_2 is considered, see Fig 5.9a. In terms of the two constants:

$$a = \frac{R_2^2\omega_2 - R_1^2\omega_1}{R_2^2 - R_1^2}\,, \qquad b = \frac{R_1^2 R_2^2(\omega_1 - \omega_2)}{R_2^2 - R_1^2}\,,$$

the resulting velocity in polar coordinates (r, φ) is $u_r(r, \varphi) = 0$, $u_\varphi(r, \varphi) = ar + b/r$, while in Cartesian coordinates, (x, y) with $r^2 = x^2 + y^2$, it can be written as:

$$\begin{aligned} u_1(x, y) &= -u_\varphi \frac{y}{r}\,, & \phi_1(x, y) &= 2b\frac{y}{r^2}\,, \\ u_2(x, y) &= u_\varphi \frac{x}{r}\,, & \phi_2(x, y) &= -2b\frac{x}{r^2}\,. \end{aligned}$$

The following cases are explored:

[TF.1] $R_1 = 1$, $R_2 = 2$, $\omega_1 = 1$, and $\omega_2 = 0$ with least-squares weighting omitted, that is $w_j = 1$ for $j = 1, \ldots, 4$. Linear basis functions on 3-point triangular Lagrange elements are used and pure velocity Dirichlet boundary conditions specified; in order to obtain a uniquely solvable system necessary constraints are prescribed

for the potential variables:

$$\begin{cases} u_{1,h}(p) = u_1(p) & \text{for } p \in \partial\Omega\,, \\ u_{2,h}(p) = u_2(p) & \text{for } p \in \partial\Omega\,, \\ \phi_{1,h}(p) = 0 & \text{for } p \in \{(-1,0);(2,0)\}\,, \\ \phi_{2,h}(-1,0) = 0\,. \end{cases}$$

Initially, a simply connected domain is considered which is achieved by additionally specifying the solution along an arbitrary cross section.

[TF.2] The same configuration as TF.1 but using quadratic polynomial basis functions on 6-point Lagrange triangles. Here, in addition, the curved boundary is approximated by piecewise quadratic edges.

[TF.3] The same configurations as in TF.1 and TF.2 but now considering the full Taylor-Couette problem in a multiply-connected domain.

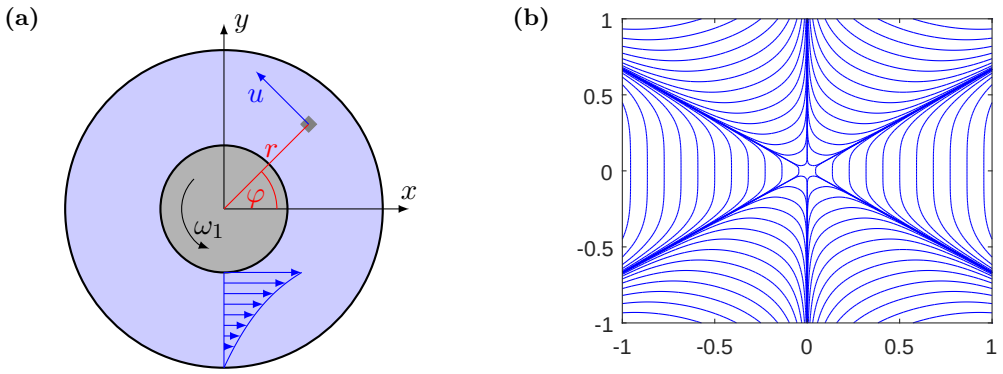


Figure 5.9: (a) Taylor-Couette flow between two concentric circles with different angular velocity [TF]. (b) Streamlines for Colliding flow [CF] in the interval $[-1, 1]$.

The results obtained for cases TF.1-TF.3 are provided in tabular form in Appendix C.1 and displayed graphically in Figures 5.10-5.12. Fig. 5.10 shows convergence results for both test case TF.1 and TF.2. Both tests were carried out with equal least-squares weighting of the four partial differential equations and calculated on approximately homogeneous increasingly refined grids, with grid width parameter h. The graphs show the global error versus the reciprocal of h, using logarithmic scales, in which according to Table C.1 the finest grid corresponds to approximately 250 thousand grid points and 1 million variables. The error is measured using three different norms: the energy norm $\|u\|_a = \sqrt{a(u,u)}$, which according to Corollary 4.10 is equivalent to the $H^1(\Omega)$ norm, the $L^2(\Omega)$ norm and the maximum norm. For linear FEs the error scales linearly, i.e. $\mathcal{O}(h)$, in

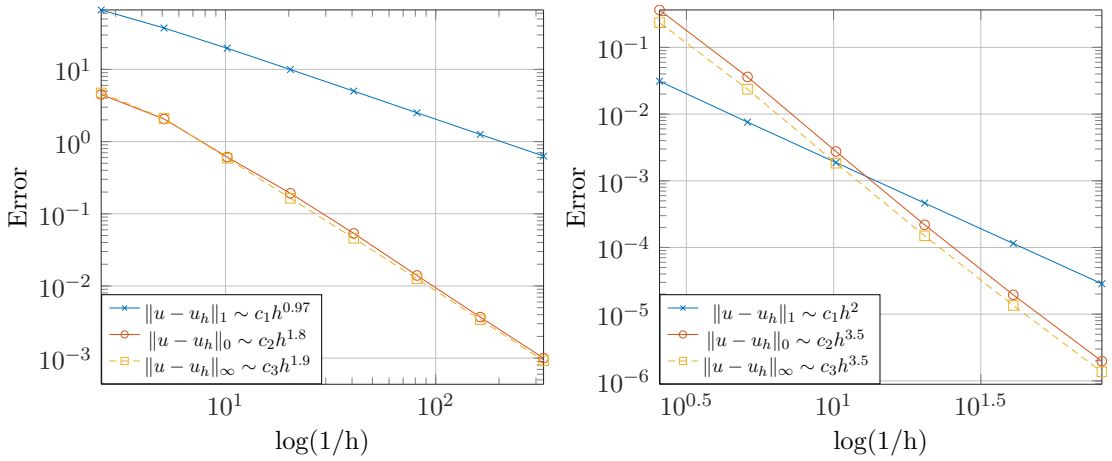


Figure 5.10: Convergence results for test case TF.1 (left) and TF.2 (right) showing global errors in terms of the energy, the $L^2(\Omega)$, and the maximum norm.

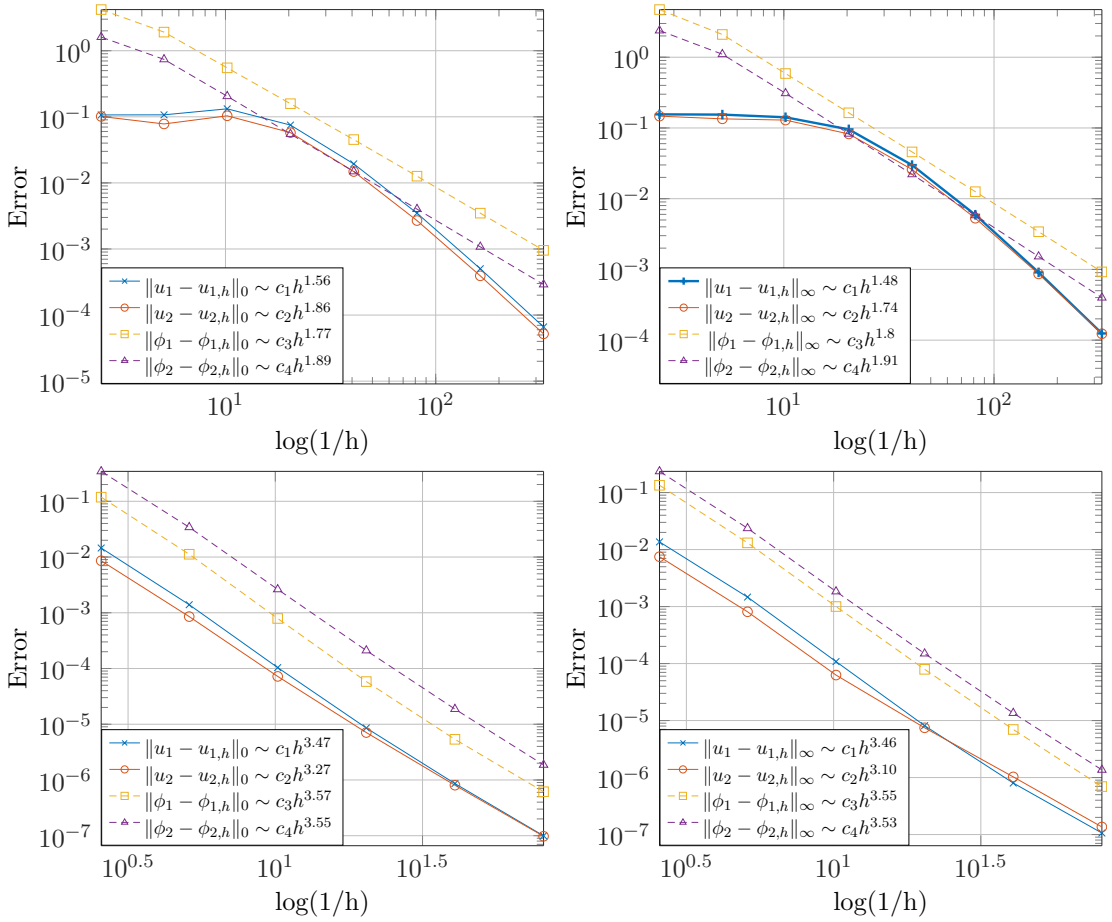


Figure 5.11: Variable-wise convergence results for test case TF.1 (top row) and TF.2 (bottom row) with errors displayed in term of the $L^2(\Omega)$ (left) and the maximum (right) norm.

the H^1-norm and quadratically, i.e. $\mathcal{O}(h^2)$, in the L^2-norm; the maximum norm behaves similar to the L^2-norm, as expected. For quadratic elements the exponents increase by one, leading to quadratic and cubic scalings, respectively. In summary, the errors behave precisely as predicted theoretically by Theorems 4.12 and 4.14, thus providing numerical verification for the optimal error theory from Chapter 4.

Fig. 5.11 shows convergence results for TF.1 and TF.2, this time for the individual variables u_1, u_2 and ϕ_1, ϕ_2. Those in the left column corresponds to L^2 error measurements and those in the right column to maximum error measurements. The scalings once again turn out to be optimal, as expected; although for the velocity components, in the case of linear FEs, is dependent on the degree of mesh refinement. The full multiply connected case, TF.3, exhibits no qualitative differences. The global error scaling displayed in Fig. 5.12 are compatible to those in Fig. 5.10.

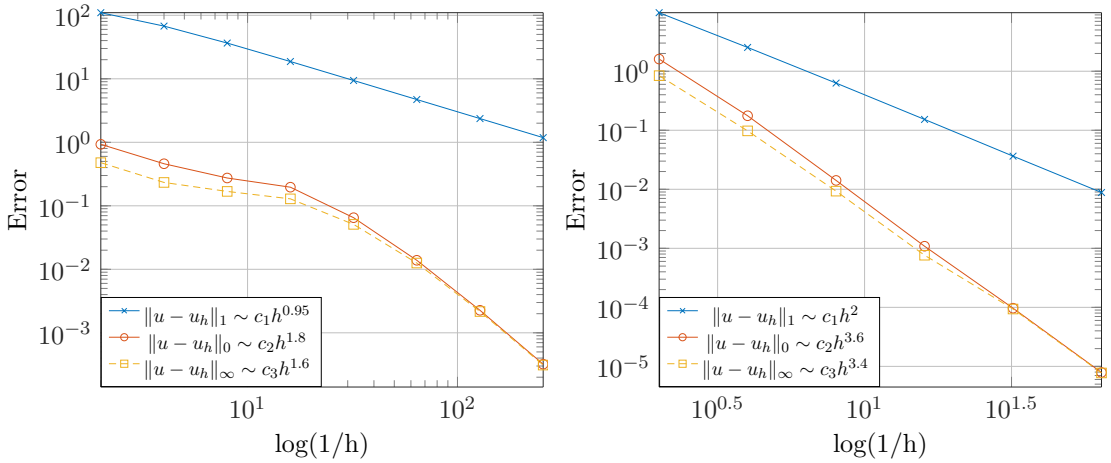


Figure 5.12: Convergence results for test case TF.3 using linear (left) and quadratic (right) basis functions, showing global errors in terms of the energy, $L^2(\Omega)$, and the maximum norm.

(b) Colliding flow [CF]

Colliding Stokes flow with an exterior force $f = 0$ is considered on the unit square $[0,1]^2$, see Fig. 5.9b:

$$\begin{cases} u_1(x,y) = 20xy^3\,, \\ u_2(x,y) = 5(x^4 - y^4)\,, \\ \phi_1(x,y) = -20(x^3y + xy^3)\,, \\ \phi_2(x,y) = -30x^2y^2 + 15y^4 - 5x^4\,, \end{cases} \tag{5.78}$$

when either pure velocity or potential Dirichlet boundary conditions or three variants of mixed conditions are prescribed:

[CF.1] Quadratic basis functions and 6-point triangular Lagrange FEs are used with no least-squares weighting applied to the equations. Pure velocity Dirichlet boundary conditions are specified and in order to obtain a uniquely solvable system the following constraints are prescribed for the potential variables:

$$\begin{cases} u_{1,h}(p) = u_1(p) & \text{for } p \in \partial\Omega\,, \\ u_{2,h}(p) = u_2(p) & \text{for } p \in \partial\Omega\,, \\ \phi_{1,h}(p) = 0 & \text{for } p \in \{(0,0);(1,1)\}\,, \\ \phi_{2,h}(0,0) = 0\,. \end{cases}$$

The case of pure potential Dirichlet boundary conditions can be handled analogously leading to nearly identical results and is therefore omitted.

[CF.2] Using the same quadratic basis functions as in CF.1, a combination of velocity and potential Dirichlet boundary conditions are prescribed such that for $x \in [0,1]$, $y \in (0,1)$:

$$\begin{cases} u_{1,h}(x,0) = u_1(x,0)\,, & u_{1,h}(0,y) = u_1(0,y) \\ u_{2,h}(x,0) = u_2(x,0)\,, & u_{2,h}(0,y) = u_2(0,y) \\ \phi_{1,h}(x,1) = \phi_1(x,1)\,, & \phi_{1,h}(1,y) = \phi_1(1,y) \\ \phi_{2,h}(x,1) = \phi_2(x,1)\,, & \phi_{2,h}(1,y) = \phi_2(1,y)\,. \end{cases}$$

[CF.3] Using the same quadratic basis functions as in CF.1 and CF.2, a combination of velocity and potential Dirichlet boundary conditions are prescribed such that for $x \in [0,1]$, $y \in (0,1)$:

$$\begin{cases} u_{1,h}(x,0) = u_1(x,0)\,, & u_{1,h}(x,1) = u_1(x,1)\,, \\ u_{2,h}(x,0) = u_2(x,0)\,, & u_{2,h}(x,1) = u_2(x,1)\,, \\ \phi_{1,h}(0,y) = \Phi_1(0,y)\,, & \phi_{1,h}(1,y) = \phi_1(1,y)\,, \\ \phi_{2,h}(0,y) = \Phi_2(0,y)\,, & \phi_{2,h}(1,y) = \phi_2(1,y)\,. \end{cases}$$

[CF.4] Using the same quadratic basis functions as above, a combination of velocity and potential Dirichlet boundary conditions are prescribed such that for $x \in [0,1]$, $y \in (0,1)$:

$$\begin{cases} u_{2,h}(x,0) = u_2(x,0)\,, & u_{2,h}(x,1) = u_2(x,1)\,, \\ \phi_{1,h}(x,0) = \phi_1(x,0)\,, & \phi_{1,h}(x,1) = \phi_1(x,1)\,, \\ u_{1,h}(0,y) = u_1(0,y)\,, & u_{1,h}(1,y) = u_1(1,y)\,, \\ \phi_{2,h}(0,y) = \phi_2(0,y)\,, & \phi_{2,h}(1,y) = \phi_2(1,y)\,. \end{cases}$$

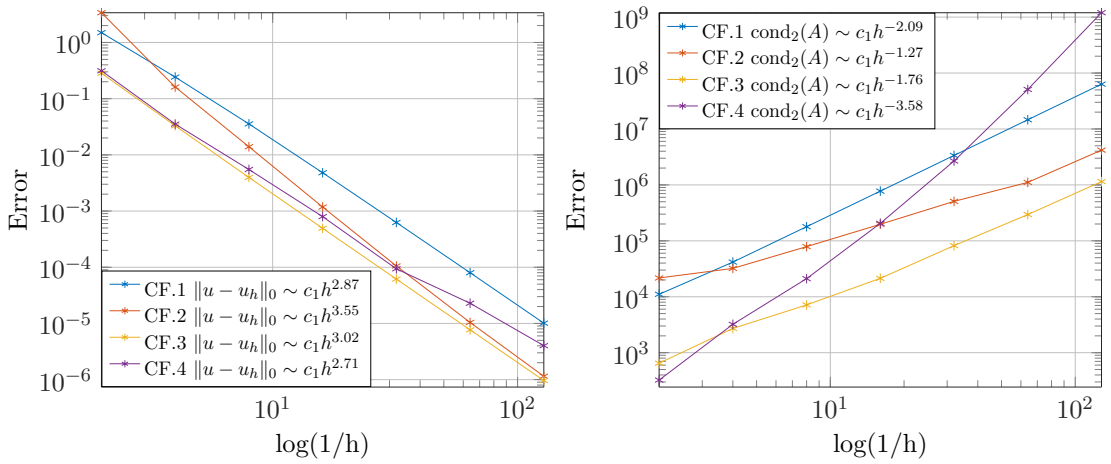


Figure 5.13: Convergence results for CF.1 - CF.4 solved on an unstructured grid (left); the discrete 2-norm condition numbers of the resulting system matrices (right).

The impact of different types of boundary conditions on the convergence behaviour and also on the condition number of the system matrices is explored. In contrast to the Taylor-Couette flow the present set-up exhibits additional complications imposed by the four irregular boundary corner points. These points not only exhibit sharp geometric corners with 90° angles, they are also contact points for different boundary conditions: In case CF.1 there are no contact points since only velocity conditions are prescribed. Cases CF.2 and CF.3 exhibit 2 and 4 contact points, respectively, but on each edge either both velocity components or both potential components are prescribed; such a combination of boundary conditions is, for instance, typical for free surface problems as described in Sec. 5.1.2 with the difference that in the example of periodic film flow the free surface does not come into contact with the rigid wall. The last case CF.4 is the most complex, since here again 4 contact points occur and on each of the edges *different combinations* of u and ϕ components are prescribed; such boundary conditions may be physically relevant when an outflow is present, as in the case of the backward-facing step problem investigated below.

Once again a sequence of homogeneous and increasingly refined grids with mesh width parameters h is utilised, with corresponding results provided in tabular form in Appendix C.2 and displayed graphically in Fig. 5.13. The left-hand plot demonstrates that optimal error reduction, proportional to h^3, is achieved in all four test cases independent of the imposition of irregular points. This is not unexpected as the convergence rates from Theorem 4.14 are in principle valid also for irregular domains as long as a $H^3(\Omega)$ solution can be guaranteed (when using 2nd-order polynomials) which is obviously true as prescribed by (5.78).

The right-hand plot of Fig. 5.13 displays the 2-norm condition number of the resulting linear system versus the reciprocal of the mesh width parameter h, both using a logarithmic scale. The condition numbers for cases CF.1-CF.3 are obviously bounded by ch^{-2} and thus verify the considerations of Sec. 5.2; however, the case CF.4 with mixed component boundary conditions appears to scale with h^{-4} indicating the linear system is significantly more difficult to solve with standard iterative solvers.

(c) Flow past a backward-facing step [BS]

In order to investigate the issue of mass conservation for the least-squares first integral method mentioned in Sec. 5.1.1, the benchmark test case of flow past a backward-facing step, see Fig. 5.14, is explored; the lower and upper walls are rigid, while on the left-hand side there is an inflow prescribed in the form of a Poiseuille profile. The outflow boundary conditions on the right is handled naturally by prescription of the components u_2 and ϕ_2 as defined in the following boundary value problem:

$$\begin{cases} u_{1,h}(0,y) = 2y(1-y) & \text{for } y \in (0,1)\,, \\ u_{1,h}(p) = 0 & \text{for } p \in \{(x,y) \in \partial\Omega \,|\, 0 < x < 5\}\,, \\ u_{2,h}(p) = 0 & \text{for } p \in \partial\Omega\,, \\ \phi_{1,h}(5,-1) = 0\,, & \\ \phi_{2,h}(5,y) = 0 & \text{for } y \in (-1,1)\,. \end{cases}$$

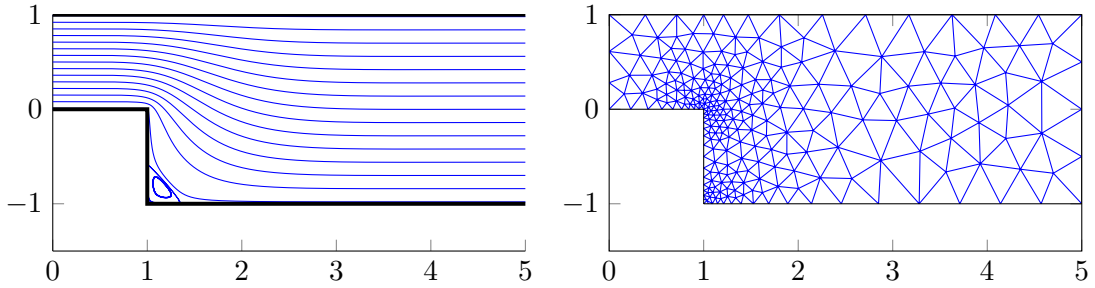


Figure 5.14: Flow geometry and streamlines (left) for the backward-facing step problem and a typical unstructured triangular grid with local refinement at the relevant corners (right).

The following alternatives are considered, including various combinations of them: (1) a homogeneous grid structure with a constant mesh width parameter or a locally refined mesh, as illustrated in Fig. 5.14, is employed, both have approximately the same number of grid points; (2) two weighting variants are tested, either $w = (1,1,1,1)$ corresponding to an equal weighting or $w = (1,1,100,1)$ corresponding to an overweight of the continuity

equation by a factor of 100; (3) as in the earlier studies both linear and quadratic elements are tested.

Fig. 5.15 shows the loss in volume flow compared to the specified inflow measured at a finite number of 50 equidistant cross sections $x = c \in [0, 5]$, with the maximum value being identified, termed the maximum loss of volume flow and measured as a percentage. The maximum loss is calculated for increasingly refined grids and displayed versus the number of grid points n_p using logarithmic scales. The left and right-hand plots of Fig. 5.15 show results for linear and quadratic elements; for both element types the two grid types and weighting variants lead to four possible scenarios.

It is observed that very coarse grids in combination with linear elements result in a massive and unacceptable loss of volume flow rendering the solutions effectively useless. Local mesh refinement leads to a slight improvement only, while an overweight of the continuity equation shows a significant improvement. Of course, with increasingly refined grids the FE approximation converges to the exact solution but requires in excess of 100 grid points to push the mass loss below 20% which is still unacceptable.

The situation changes with quadratic elements; the right-hand plot of Fig. 5.15 demonstrates that a combination of higher order elements (at least quadratic), strong weighting and local mesh refinement is sufficient to control the mass loss problem to an acceptable degree. The huge difference between the worst case, unweighted linear, and the best case, weighted quadratic, is exemplified in Table 5.1 for different numbers of grid points. In the worst case the mass loss for a mesh with 149/150 grid points is about 78% and in the best case only 0.17%. However, as is often the case a strong weighting of one of the equations comes at a price, namely that the positive definite least-squares problem is shifted towards a constrained (saddle point) problem identified by higher condition numbers. This is why weighting has to be applied with care and requires experience; in the present investigation weighting factors between 10^2 and 10^4 proved optimal.

n_p	linear, $w = 0$	n_p	quadratic, $w = 1000$
149	−77.83	150	−0.1662
332	−63.66	331	−0.1032
851	−43.16	855	−0.0223
4167	−17.86	4116	−0.0105
23323	−6.87	23891	−0.0067

Table 5.1: Maximum loss of volume flow as a function of number of grid points: the worst and best cases being linear FEs on a homogeneous grid without least-squares weighting and quadratic FEs on a locally refined mesh with weighting, $w = 1000$, respectively.

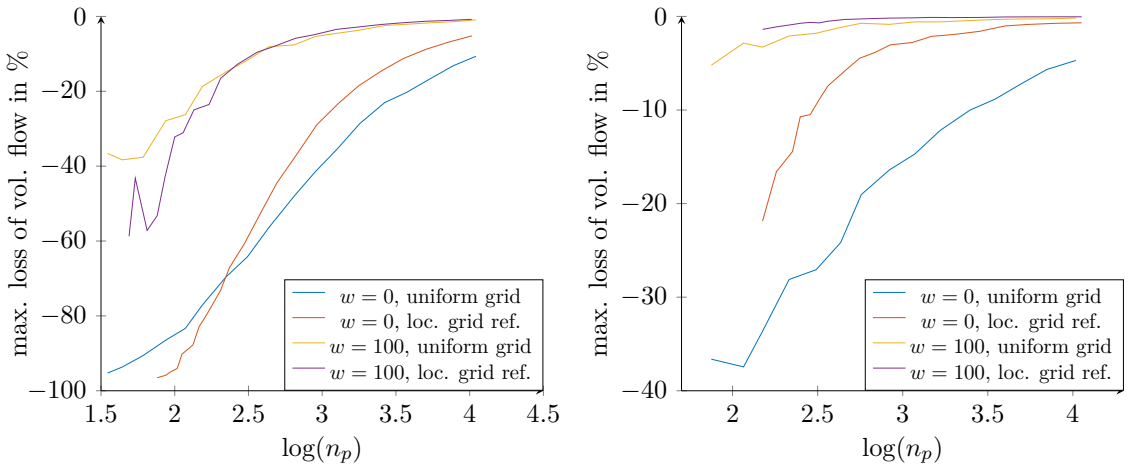


Figure 5.15: Maximum loss in volume flow measured over all possible cross sections and for increasingly refined grids; in the case of linear FEs (left) and quadratic FEs (right).

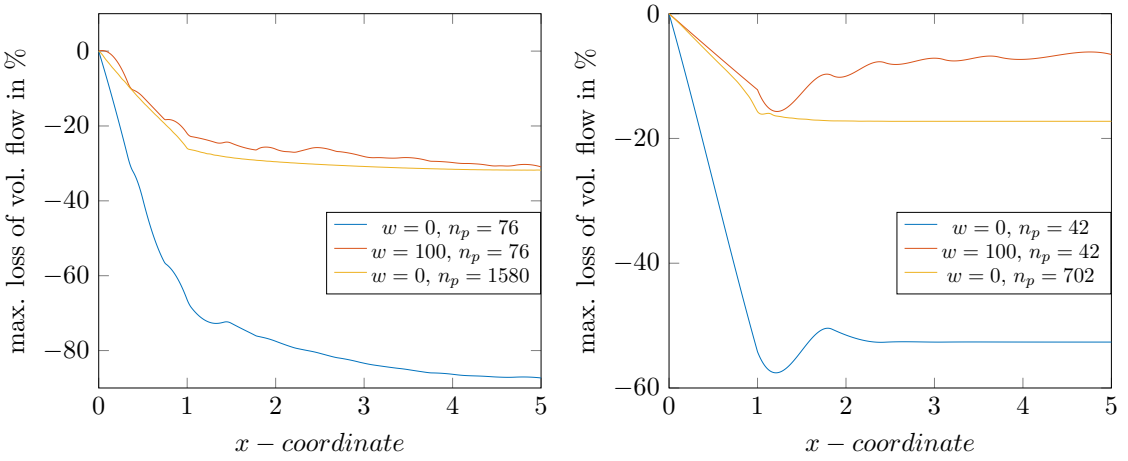


Figure 5.16: The loss of volume flow calculated for each cross-section along the x-axis for $x \in (0, 5)$ and for the cases specified: with linear FEs (left) and quadratic FEs (right).

Finally Fig. 5.16 shows the mass loss continuously distributed over all cross sections on the x-axis, for linear and quadratic elements in the left and right-hand plot, respectively. In the linear case the mass loss grows continuously with distance from the inflow but this is not the case when quadratic elements are used and in which case the maximum loss occurs just to the right of the step down. In both cases three variants are displayed: firstly, for a relatively coarse grid without weighting; secondly, for a coarse grid with a weighting factor of 100 for the continuity equation; thirdly, in the case of a significantly refined mesh without weighting. This demonstrates, for example in the quadratic element case, that without weighting about 700 grid points are necessary to achieve the same accuracy of mass conservation which, with weighting, is already achieved with only 40 grid points.

5.3.2 Non-linear test cases and results

As a last step and for completeness the performance of the full non-linear NS solver is investigated using the well-known lid-driven cavity problem as a test case, solved on the unit square $[0, 1]^2$ with the usual velocity boundary conditions as indicated in the top-left plot of Fig. 5.17; the streamlines shown correspond to the symmetric Stokes flow case, $\mathrm{Re} = 0$.

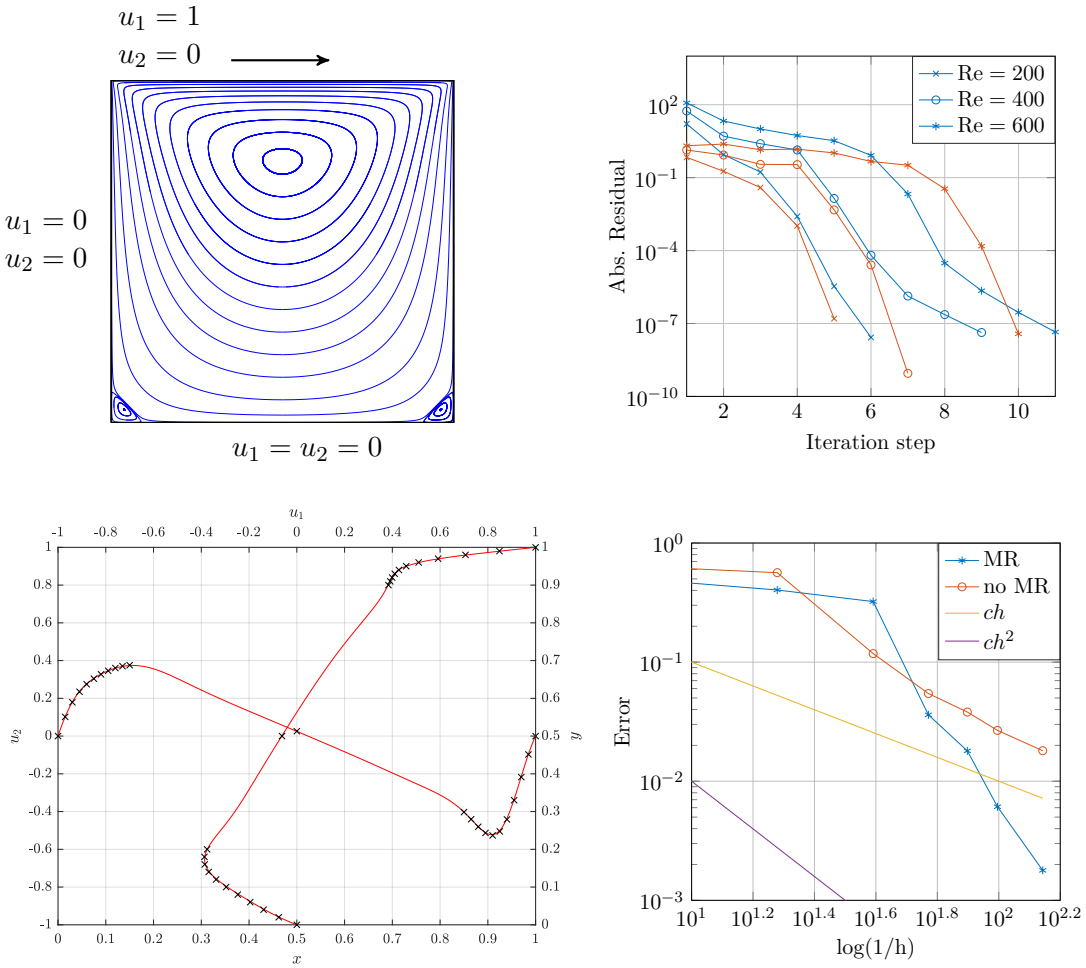


Figure 5.17: Problem set-up and streamlines ($\mathrm{Re} = 0$) for the 2D Lid-driven cavity problem (top-left); convergence of the Newton method for the LSFEM (blue) and Taylor-Hood mixed Galerkin FEM (red) for three different Reynolds numbers (top-right); u_1-velocity along the axis $x = 1/2$ and u_2-velocity along the axis $y = 1/2$ for $\mathrm{Re} = 1000$ (bottom-left) – the red lines are calculated with the LSFEM and compared to the black crosses obtained by Erturk et al. [90]; the maximum error displayed against the reciprocal of the mesh width parameter h using logarithmic scales (bottom right) – the convergence curves can be compared to the straight lines which are proportional to h and h^2.

In the top right-hand plot of Fig. 5.17 the decay of the absolute residuals in the course of the Newton iteration are shown for three different Reynolds numbers, Re = 200, 400 and 600. The blue lines correspond to the LSFEM and the red lines to a Taylor-Hood discretisation of the primitive variable formulation, see Sec. C.3 of Appendix C. The calculations were performed using an unstructured, corner-refined grid comprised of approximately 30,000 nodes and 15,000 triangular elements for the quadratic basis functions. Here and in subsequent calculations a least-squares weighting was chosen with $w = [1,\, 1,\, 10^4,\, 10^4]$. Both approaches lead to a comparable number of Newton iterations to reduce the absolute residual beyond a specified threshold of 10^{-7}. It is observed that the convergence behaviour of the mixed FEM is somewhat more uniform: after starting off slowly the length of which increases with increasing Re, a phase of fast quadratic convergence is reached. In contrast, the LSFEM associated with the first integral reaches a subsequent third phase of slower convergence towards the end; showing that the reduced non-linearity of the first integral formulation does not exhibit any advantage over a standard one for moderate Reynolds numbers of a few hundred. However, it should be noted that the radius of convergence of the LSFEM is slightly larger: Even for Re = 600 it can be seen that the latter reaches a point of faster convergence earlier and this trend is reinforced for higher Reynolds numbers, delaying – compared to the mixed FEM – the onset of divergence. The standard remedy for this for highly convection dominated flows is to employ a Reynolds number stepping methodology.

Consider now the bottom-left plot in Fig. 5.17 which displays the u_1-velocity along the axis $x = 1/2$ and the u_2-velocity along the axis $y = 1/2$ for the case Re = 1000. The red lines are calculated with the LSFEM and a grid width parameter of $h = 1/100$ with additional corner mesh refinement leading to approximately 36,000 elements and 72,000 nodes. These are compared with the black crosses obtained by Erturk et al. [90] using a high-order FD scheme on a uniform regular grid containing 601×601 points. The two are seen to be in very good qualitative agreement. This agreement is captured in more detail in the bottom-right plot of Fig. 5.17, where the error is displayed versus the reciprocal of the mesh width parameter h, using logarithmic scales, for which the number of grid points varies between approximately 3700 and 161,000. The error in the maximum norm is approximated by taking the maximum absolute deviation from the solution by Erturk et al. [90]. Two cases are shown: the convergence of LSFEM with and without local mesh refinement (MR) at the corners; the convergence curves can be compared to the straight lines which have been included and are proportional to h and h^2. Obviously convergence without local MR starts approximately quadratically and approaches a linear asymptote only, while convergence with local MR is able to restore the optimal cubic convergence after starting off slowly. Note, that the maximum norm also represents an upper bound for the L^2-norm.

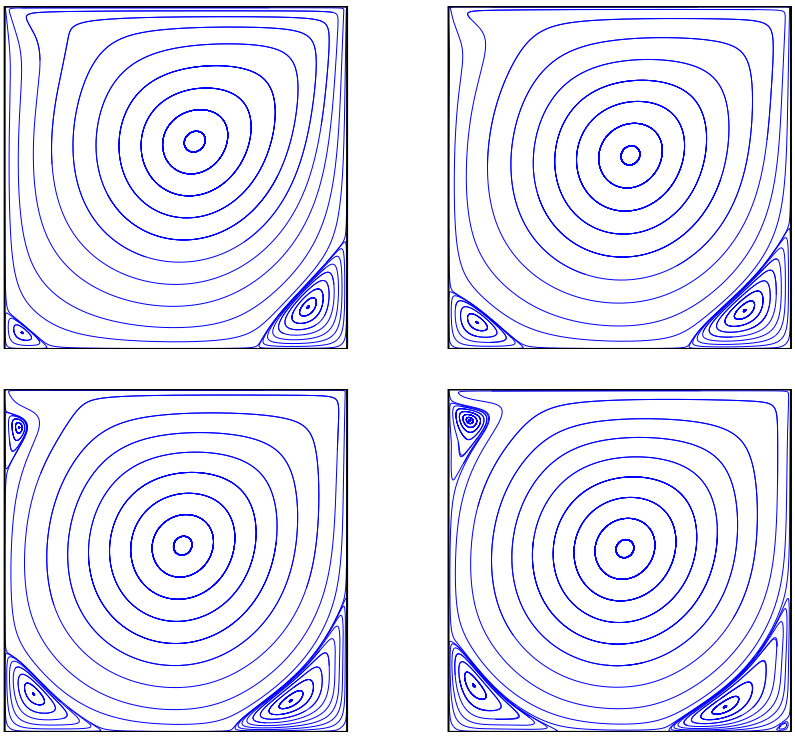


Figure 5.18: Streamline plots for the 2D Lid-driven cavity problem and Reynolds numbers Re = 400 (top-left), 1000 (top-right), 2500 (bottom-left), 5000 (bottom-right) obtained via the LSFEM.

A nearly optimal convergence rate for the lid-driven cavity problem is rather unexpected in the light of the results of Sec. 4.3.4. The 90°-corners at the bottom of the domain pose no problems as the boundary data is otherwise smooth; here, the solution is even locally in H^3, cf. Fig. 4.1, which is the necessary requirement for optimal convergence of second-order Lagrange elements. However, the u_1 velocity component exhibits discontinuities at both upper corners locally leading to a further decrease of regularity. Recapitulating the considerations in Appendix B.4, the transcendental equation (B.77) characterises the solutions of the homogeneous boundary-value problem, but in the present example the singularity is to leading order determined by the particular solution of the non-homogeneous problem in the vicinity of the corner. Corresponding calculations of the particular Stokes solution, which are in accord with Luchini [166], result in $u_1(r,\varphi) = f(\varphi)$ in polar coordinates only depending on the angle but not on the radius. In Cartesian coordinates this leads to a gradient $\nabla_{(x,y)} u_1 \notin L^2$ which goes to infinity and is not square-integrable; thus the local solution is characterised by $u \in H^1_\beta \not\subset H^1$ for some $\beta > 0$. Although the present H^1 discretisation is non-conforming in this respect – unless modifications are performed as in [136, 157] – a strong mesh refinement seems to be able to recover more or less the complete convergence loss.

For illustration purposes Fig. 5.18 shows streamlines for the 2D Lid-driven cavity problem for Reynolds numbers of Re = 400, 1000, 2500, 5000. The computations were performed with least-squares weighting and Newton iteration and stopped when the corresponding residual reached a threshold of 10^{-8}. The LSFEM was employed with quadratic basis functions and a corner-refined unstructured grid having approximately 200,000 nodes and correspondingly 100,000 elements; this would be equivalent to a uniform 450×450 mesh. The resulting linear systems involved roughly 800,000 unknowns and a matrix with approximately 35 million non-zero entries. The streamline plots shown are in excellent agreement with corresponding investigations in the literature, see [90].

5.4 Algebraic multigrid

Multigrid methods, geometric (GMG) and algebraic (AMG), essentially rely on two basic principles: smoothing and coarse-grid correction [119, 259]. The former exploits the observation that classical relaxation techniques, for instance Jacobi or Gauss-Seidel (GS) procedures, exert a strong smoothing effect on the error, efficiently reducing high frequency error components. While low frequency error components are reduced much more slowly, according to the coarse-grid principle they can be well approximated on coarser grids; here, fine grid low frequency errors become higher frequency ones and can be more easily reduced by classical relaxation methods. By a suitable combination of smoothing and coarse-grid correction a recursively applicable scheme is achieved capable of efficiently reducing error components of all frequencies. In principle by smart tuning of the multigrid parameters, an asymptotically optimal method is possible allowing for a computing effort which, in contrast to most conventional methods, increases linearly with the number of unknowns.

In contrast to GMG methods, AMG only utilises information given by the system matrix and consequently is less problem specific and applicable in a broader context. AMG is particularly well suited to problems involving complex and unstructured grid systems. Certainly, the spirit behind this development is the desire to provide multigrid performance in an intelligent black box like setting without any need for problem-dependent adaptation and tuning; an objective far from being realised at the present time. AMG utilises the fact that all information about the grid and the variable coupling is essentially included in the system matrix. Thus a matrix-based method can be constructed that is not founded on grids, partial grids and nodes but on sets of variables, subsets of these and single variables. Without knowledge of the geometry the system is interpreted as an adjacency matrix, apropos graph/network theory, and the variables as its nodes. Against this background a recursive scheme of smoothing and coarse-grid correction is established

similar to GMG. AMG was introduced in the 1980s for the solution of scalar elliptic PDEs [45, 46, 56, 210, 242, 259] and has been continuously developed resulting in variants such as aggregative AMG, smoothed aggregation (SA) [51, 242, 263], extensions to PDE systems [65] and in customised versions for FE applications (AMGe) [50, 126]. Moreover, intelligent adaptive mechanisms [51, 52, 66, 67] have been devised and efficient parallel implementations constructed [93, 152, 212].

In the present work the classical ideas of scalar elliptic AMG [45, 46, 210, 242] are adopted to solve linear systems resulting from the LSFEM described above. To what extent this is possible forms part of the investigation, i.e., how close the corresponding linear systems fit into the class of so-called M-matrices, that classical convergence analysis relies on. Clearly the desired objective is the construction of an efficient and, at the same time, flexible PDE algebraic multigrid procedure which easily deals with unstructured grids, with different types and combinations of FEs, which exploits the beneficial positive definiteness of the linear systems and at best exhibits a linearly increasing computational effort with regard to the number of variables. The AMG strategy illustrated in Sec. 5.4.1 – 5.4.2 is confined to a scalar and serial algorithm and is substantially based on the well-established monograph by Stüben [242], which also served as a starting point for the Matlab implementation in Sec. 5.4.4. Extension to PDE strategies in Sec. 5.4.3, particularly the so-called unknown-based AMG, relies on the work of Clees [65].

5.4.1 Scalar AMG: basic ideas

Before focussing on specific components of AMG a short but sufficiently detailed overview of the method is provided together with introduction of the necessary underpinning notation. For the sake of clarity the description is initially restricted to a two-level MG cycle (Sec. 5.4.1(a)) using geometrically motivated notation, i.e., the index h indicates the fine level and H the coarse level; subsequently in Sec. 5.4.1(b) the full V-cycle and its variants are specified.

(a) Two-level cycle

A large and sparse linear system of the form:

$$A_h u^h = f^h \qquad \text{or} \qquad \sum_{j \in \Omega^h} a_{ij}^h u_j^h = f_i^h \quad (i \in \Omega^h)\,, \tag{5.79}$$

is to be solved without any knowledge of the associated geometric information. Let $A_h \in \mathbb{R}^{n \times n}$ and $u_h, f_h \in \mathbb{R}^n$ denote the system matrix, solution and the right-hand side on the fine level with A_h assumed to be symmetric and positive definite; Ω^h denotes the index set $\{1, 2, ..., n\}$ of all fine level variables.

The AMG is based on the interplay between smoothing and coarse-level correction; however, in a pure algebraic context *smoothness* has to be suitably redefined without reference to grids, see Sec. C.4.1 of Appendix C. For now, assume a linear smoothing operator $S_h : \mathbb{R}^n \to \mathbb{R}^n$ to exist which smooths the error in a "reasonable" algebraic sense. Let $u_\star^h$ denote the exact solution of the linear system (5.79) on the fine level and u^h and $\bar{u}^h$ the current approximation before and after one smoothing step, respectively; the corresponding errors are $e^h = u_\star^h - u^h$ and $\bar{e}^h = u_\star^h - \bar{u}^h$. Assuming that $\bar{e}^h$ $(= S_h e^h)$, compared to e^h, is algebraically smoothed, a smoothing step $u^h \to \bar{u}^h$ with corresponding error smoothing $e^h \to \bar{e}^h$ is given by:

$$\bar{u}^h = S_h u^h + (I_h - S_h) A_h^{-1} f^h \,, \tag{5.80}$$

$$\bar{e}^h = S_h e^h \,. \tag{5.81}$$

Typical smoothers are linear iterative solvers such as the Jacobi, GS or successive over-relaxation (SOR) schemes, sometimes more expensive solvers based on a an incomplete lower-upper (ILU) decomposition are applied [119, 211, 259]; however, most of these methods can be represented by means of a defect correction of the form:

$$S_h := I_h - B_h^{-1} A_h \,, \tag{5.82}$$

in which B_h defines a reasonably good approximation to A_h such that B_h is comparatively easy to invert and the inverse B_h^{-1} remains sufficiently sparse. The simplest variants are obtained for B_h equal to the diagonal part or the lower/upper triangular part of A_h which leads to the Jacobi and the GS-smoothers. Using definition (5.82) transforms equation (5.80) into the form:

$$\bar{u}^h = u^h + B_h^{-1}(f^h - A_h u^h) \,,$$

which will be used later.

If u^h is the current approximation after smoothing with corresponding residual $r^h = f^h - A_h u^h$, the error components that could not be reduced by algebraic smoothing have to be handled via coarse-grid correction, as illustrated in Fig. 5.19. Thus the system of equations for the error, namely $A_h e^h = r^h$, is approximated on a coarser level, the construction of which requires an algorithm that splits the total index set Ω^h into two disjoint subsets, $\Omega^h = C^h \cup F^h$, in which case the sets $C^h = \Omega^H$ and F^h include the variable indices only existing on the coarse and fine levels, respectively. The coarsening procedure should follow the direction of algebraic smoothness, see Sec. C.4.2.

In the following a splitting of the index set into C- and F-variables is assumed to be given with $n_c = |C^h| = |\Omega^H| < n$ denoting the number of coarse level variables; then

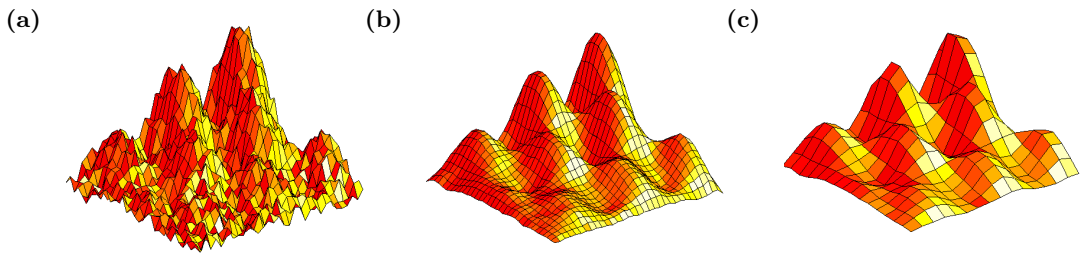


Figure 5.19: An illustration of error evolution in the course of an AMG: (a) the initial error, (b) the error after smoothing and (c) approximation of the smoothed error on a coarser grid.

appropriate transfer operators are required allowing movement between the levels, see Sec. C.4.3. $I_h^H : \mathbb{R}^n \to \mathbb{R}^{n_c}$ and $I_H^h : \mathbb{R}^{n_c} \to \mathbb{R}^n$ denote the linear restriction and interpolation operators mapping fine level vectors to coarse level vectors and vice versa; these operators can be represented by matrices $I_h^H \in \mathbb{R}^{n_c \times n}$ and $I_H^h \in \mathbb{R}^{n \times n_c}$ which are subsequently assumed to be of full column and row rank, respectively. Moreover, for symmetry reasons the restriction operator is always assumed to equal the transpose of the interpolation operator, that is $I_h^H = (I_H^h)^T$.

The linear system for the error on the coarse level will be of the form:

$$A_H e^H = r^H \qquad \text{or} \qquad \sum_{l \in \Omega^H} a_{kl}^H e_l^H = r_k^H \quad (k \in \Omega^H)$$

with $A_H \in \mathbb{R}^{n_c \times n_c}$ and $e^H, r^H \in \mathbb{R}^{n_c}$, respectively, and is constructed in accordance with the Galerkin principle [242]. The residual vector on the coarse level is obtained by restriction of the corresponding fine level vector and the Galerkin matrix A_H is constructed by means of the fine level matrix A_h and the transfer operator, i.e.,

$$\begin{aligned} r^H &:= I_h^H r^h = (I_H^h)^T r^h \,, \\ A_H &:= I_h^H A_h I_H^h = (I_H^h)^T A_h I_H^h \,. \end{aligned}$$

At the coarse-grid level the linear system can be solved with considerably less effort and the associated solution e^H interpolated back to the fine level to correct the former approximate solution u_{old}^h. A two-level correction step leading from u_{old}^h with error e_{old}^h to a new approximation u_{new}^h with error e_{new}^h can be written as:

$$u_{new}^h = u_{old}^h + I_H^h e^H \qquad \text{involving} \qquad A_H e^H = I_h^H (r_{old}^h) = I_h^H (f^h - A_h u_{old}^h) \,, \tag{5.83}$$

$$e_{new}^h = K_{h,H} e_{old}^h \qquad \text{with} \qquad K_{h,H} := I_h - I_H^h A_H^{-1} I_h^H A_h \,. \tag{5.84}$$

$K_{h,H}$ denotes the so-called coarse-grid correction operator. The error reduction of a com-

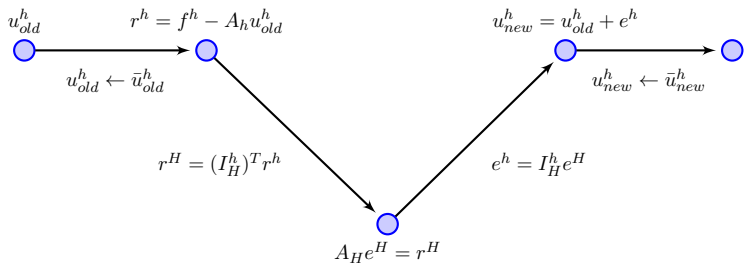


Figure 5.20: Schematic of a two-level cycle involving pre-smoothing, coarse-grid correction and post-smoothing.

plete two-level cycle consisting of ν_1 pre-smoothing steps, a coarse-grid correction step and ν_2 post-smoothing steps, as illustrated schematically in Fig. 5.20, is characterized by the following two-level iteration operator $M_{h,H}$:

$$e^h_{new} = M_{h,H} e^h_{old} \qquad \text{with} \qquad M_{h,H}(\nu_1, \nu_2) = S_h^{\nu_2} K_{h,H} S_h^{\nu_1} .$$

(b) V-cycle and more complex variants

When applied recursively the above gives rise to the so-called V-cycle, Algorithm 3, which is comprised essentially of a setup (Algorithm 4) and a cycle (Algorithm 5) phase. In the setup phase, based on the initial system matrix A, recursively coarsened levels $k = 1, 2, \ldots, M$ are defined with corresponding variable index sets $\Omega^1 \supset \Omega^2 \supset \cdots \supset \Omega^M$ and corresponding sets for coarse and fine level variables $\Omega^k = C^k \cup F^k, k = 1, 2, \ldots, M-1$. If the maximum number of variables of the linear system falls below a pre-estimated value, the coarsening process stops. For each new level $k = 2, 3, \ldots, M$ the interpolation matrix I_k^{k-1} and the Galerkin operator A_k are calculated and additionally a smoothing operator S_k is determined. All this information is calculated and saved only once and then iteratively utilised in the cycle phase.

Algorithm 3 Classical AMG method

```
1: procedure u = AMG(A, f, u, n)
2:     SETUP(A, n).
3:     Set r := f − Au.
4:     while ‖r‖ > ϵ do
5:         e := CYCLING(r, 1).
6:         Set u := u + e.
7:     end while
8:     return u.
9: end procedure
```

Algorithm 4 AMG Setup

procedure SETUP(A, n)
  Set $k := 1$.
  Set $A_k := A$.
  Set $\Omega^k := \{1, 2, \ldots, n\}$.
  repeat
    $[C^k, F^k]$:= SPLITTING(A_k, Ω^k).
    (Split Ω^k into two disjoint subsets $\Omega^k = C^k \cup F^k$).
    Set $\Omega^{k+1} := C^k$.
    I_{k+1}^k := INTERPOL(A_k, C^k, F^k).
    (Compute interpolation matrix I_{k+1}^k).
    Compute Galerkin operator $A_{k+1} := (I_{k+1}^k)^T A_k I_{k+1}^k$.
    Define smoothing operator S_k involving the approximation $B_k \approx A_k$.
    Set $k := k + 1$.
  until $|\Omega^k|$ small enough.
  Set $M := k$.
  Save setup data in module DATA:
  I_{i+1}^i $(i = 1, \ldots, M-1)$, A_i $(i = 1, \ldots, M)$, M.
end procedure

Algorithm 5 AMG V-cycle

procedure u^k = CYCLING(f^k, k)
  Load module DATA.
  if $k = M$ **then**
    Solve $A_M u^M = f^M$ by a direct solver.
  else
    Set $u^k := B_k^{-1} f^k$.
    for $i = 1, \ldots, \nu_1$ **do**
      Set $u^k := u^k + B_k^{-1}(f^k - A_k u^k)$.
    end for
    (Apply ν_1 times the smoothing operator $S_k := I_k - B_k^{-1} A_k$).
    Execute coarse level correction:
    Set $r^{k+1} := (I_{k+1}^k)^T (f^k - A_k u^k)$.
    e^{k+1} := CYCLING(r^{k+1}, $k+1$).
    Interpolate and correct $u^k := u^k + I_{k+1}^k e^{k+1}$.
    for $i = 1, \ldots, \nu_2$ **do**
      Set $u^k := u^k + B_k^{-T}(f^k - A_k u^k)$.
    end for
    (Apply ν_2 times the smoothing operator $S_k := I_k - B_k^{-T} A_k$).
  end if
  return u^k.
end procedure

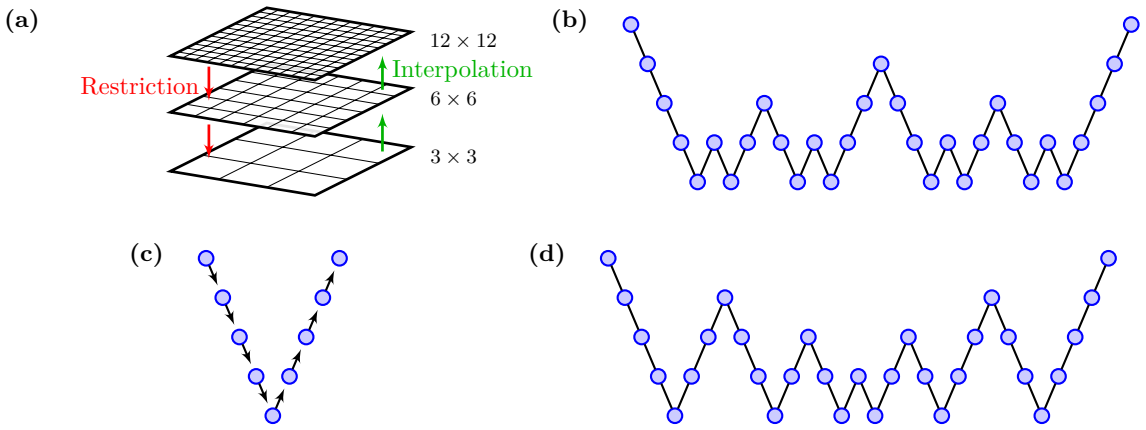


Figure 5.21: Schematic of different multigrid cycling strategies: (a) illustrates restriction (red arrows) and interpolation (green arrows) between three different grid hierarchies using structured Cartesian refinement; (b) recursive W-cycle; (c) V-cycle; (d) a symmetrised full multigrid cycle (F-cycle), half of which forms the classical full multigrid cycle.

The cycling process continues until the residual norm reaches a specified threshold value ϵ. Every cycle is initialised by a right-hand side f, and the level index k; the remaining data is obtained from the setup phase. One cycle consists of ν_1 pre-smoothing steps, a coarse-grid correction step and ν_2 post-smoothing steps in which the smoothing, according to lines 8 and 16 of Algorithm 5, is performed in a symmetric manner. As a coarse-grid solver the cycle procedure is called until the coarsest level M is reached at which point the system is solved directly by Gaussian elimination. The computed error is interpolated onto the finest level and used there to correct the approximate solution.

A detailed descritpion of the above essential multigrid components is provided in Appendix C.4. In brief, Sec. C.4.1 considers algebraic smoothing followd by the setup phase for an appropriate coarsening procedure and a construction formula for the interpolation operators in Sec. C.4.2 and C.4.3; these components have to be chosen in a way that guarantees an efficient interplay between smoothing and coarse-grid correction and thus guarantee convergence of the iteration. Furthermore, splitting and transfer operators should be chosen such that the dimension of the Galerkin operator is sufficiently reduced while preserving a reasonable sparse structure. The convergence of the V-cycle can be improved by more elaborate and simultaneously expensive variants such as W or F type cycles [242], cf. Fig. 5.21, which normally do not pay off for pure scalar elliptic problems, but may pay off in the PDE case. In contrast the “acceleration” of AMG by an iterative solver or, in other words, the preconditioning of such an iterative solver by AMG, is nearly always worthwhile.

5.4.2 AMG as a preconditioner

In order to improve robustness, multigrid methods are rarely used as stand-alone solvers but in practice are frequently combined with so-called accelerators, namely standard iterative Krylov solvers such as the conjugate gradient (CG) method, its biconjugate stabilised variant (BiCGstab) or the generalised minimal residual method (GMRES) [211, 242]. This development was prompted by the experience that an acceleration approach is both simpler and more efficient compared to the elaborate optimisation of the numerous multigrid components.

This is also true for AMG. Research has demonstrated that AMG works as an excellent preconditioner, much better than for instance one-level methods of an ILU type; the main reason being that AMG, in contrast to one-level methods, efficiently acts on all error components, the high as well as the low frequency error components. Nevertheless, it can be beneficial to limit the cost of the setup phase in favour of appropriate acceleration. Of course one wants to "transport" the most important information between levels by optimising coarsening and interpolation but this will always remain a heuristic process and it is commonly found that error reduction efficiency is compromised for a few components of the error. In fact, a few eigenvalues of the AMG iteration matrix may be located significantly closer to 1 than the rest and consequently the convergence factor will be limited by the slow convergence of these specific components. Such components can be quickly reduced by an acceleration procedure, an exemplar of which is preconditioned CG [43, 237].

Denote the input error and the error after the i-th iteration step by e_0 and e_i; furthermore let $\sigma(A)$ denote the set of all eigenvalues of a matrix A and $\|.\|_1$ the energy norm. Define the set of all polynomials of order i by P_i, with $P_i(0) = 1$. Then under the assumption of a symmetric positive definite system matrix the CG method admits the following error estimate in the i-th iteration step [43]:

$$\|e_i\|_1^2 \leq \min_{P_i} \max_{\lambda \in \sigma(A)} [P_i(\lambda)]^2 \|e_0\|_1^2 \,.$$

As a consequence, CG solves a system with only k different eigenvalues in k iteration steps and in general the convergence speed improves for systems for which eigenvalues are strongly concentrated in few areas in comparison to systems which exhibit a homogeneous distribution of eigenvalues in the interval $[\lambda_{min}, \lambda_{max}]$. Normally, AMG methods efficiently reduce most of the error components apart from a few corresponding to algebraically smooth eigenvectors; thus the eigenvalues of the preconditioned system are strongly concentrated around 1 with few very small eigenvalues remaining near zero. In the case of n_c remaining eigenvalue clusters, the error associated with a CG-accelerated

AMG method is extremely small after just n_c iteration steps. Algorithm 6 illustrates the general procedure forming a CG-accelerated AMG.

Algorithm 6 AMG preconditioned CG method

```
1: procedure u = PCG(A, f, u, n)
2:     SETUP(A, n).
3:     r := f − Au.
4:     d = h := CYCLING(r, 1).
5:     while ‖r‖ > ϵ do
6:         z := Ad.
7:         α := r^T h/(d^T z).
8:         u := u + αd.
9:         β := r^T h.
10:        r := r − αz.
11:        h := CYCLING(r, 1).
12:        β := r^T h/β.
13:        d := h + βd.
14:    end while
15: end procedure
```

5.4.3 AMG and the first integral approach

Above and in Sec. C.4 of Appendix C, the foundations of an algebraic multigrid solver for scalar elliptic differential equations have been laid which can, to a certain degree, be extended to the *systems* of PDEs [65] of interest in this thesis. Indeed, such an extension is needed to deal with the discrete systems stemming from the first integral LSFEM described earlier.

Essentially there are three different approaches for PDE systems: variable-based; unknown-based and point-based. The first two only are considered here. The first just ignores the fact that the system matrix A arises from the discretisation of a system of PDEs; instead, the AMG setup as described above is applied "as-is" to the matrix A. This so-called variable-based AMG ([65], Section 3.2) will only work well if A is a M-matrix or an essentially positive matrix, see Sec. C.4.1(b). In practice, for systems of PDEs this requires that the couplings between the different physical unknowns are very weak which is not fulfilled for the discrete systems arising from Sec. 5.1. Indeed, a series of tests using scalar AMG as a standalone solver applied to the system (5.38) for a 2D lid-driven cavity problem show, at best, extremely slow convergence. In contrast, due to the positive definiteness apropos linear systems the classical iterative CG method is more applicable and should be used to accelerate the multigrid procedure as described in Sec. 5.4.2. But the multigrid cycle is very inefficient so that the accelerated AMG is only a

slight improvement against the bare CG iteration.

In unknown-based AMG ([65], Section 3.3), the second approach for PDEs, scalar AMG algorithms are applied to each physical unknown u_i *separately*. To use this method, a variable-to-unknown mapping (VU mapping) has to be provided. Having a discrete vector $u \in \mathbb{R}^n$, this mapping identifies, for each entry u_i, the physical unknown $j \in \{1, \dots, M\}$ that u_i belongs to. In other words, the index set $\Omega = \Omega_{[1]} \cup \dots \cup \Omega_{[M]}$ is disjointly divided, where each $\Omega_{[i]}$ contains the indices corresponding to discretisation of the physical unknown u_i. Now, the matrix is reordered such that it is sorted in terms of the physical unknowns, and the block structure,

$$A = \begin{pmatrix} A_{[1,1]} & A_{[1,2]} & \cdots & A_{[1,M]} \\ A_{[2,1]} & A_{[2,2]} & \cdots & A_{[2,M]} \\ \vdots & \vdots & \ddots & \vdots \\ A_{[M,1]} & A_{[M,2]} & \cdots & A_{[M,M]} \end{pmatrix},$$

is obtained where each matrix block $A_{[i,j]}$, $i, j = 1, \dots, M$ describes the couplings between the physical unknowns i and j. The matrix blocks $A_{[i,j]}$, $i \neq j$ are not necessarily square, as the discretization meshes $\Omega_{[i]}$ for different physical unknowns may not have the same size. Each diagonal block $A_{[i,i]}$ can be viewed as the discretization of a scalar equation for the i-th unknown. In unknown-based AMG, the method is built around these scalar blocks; more precisely, for each $i = 1, \dots, M$, the following steps are performed:

1. Extraction of a strength matrix $S_{[i]}$ from the matrix entries of $A_{[i,i]}$.
2. C/F-splitting $C_{[i]} \cup F_{[i]} = \Omega_{[i]}$ based on $S_{[i]}$.
3. Building an interpolation operator $I_{[i]} : \mathbb{R}^{|C_{[i]}|} \to \mathbb{R}^{|\Omega_{[i]}|}$, and a restriction operator $R_{[i]} = I_{[i]}^T$.

Now, the different coarse-grids related to the unknowns are merged into a global coarse-grid:

$$C = \bigcup_{i=1}^{M} C_{[i]},$$

and the global interpolation (the so-called *multiple-unknown (MU)* interpolation) operator and the restriction operator are assembled according to:

$$I = \begin{pmatrix} I_{[1]} & 0 & \cdots & 0 \\ 0 & I_{[2]} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I_{[M]} \end{pmatrix}.$$

and $R = I^T$. Finally the coarse-grid operator is assembled via the full Galerkin product as:

$$A^C = I^T A I \,. \tag{5.85}$$

A different way to define the coarse-grid operator is to employ the diagonal block entries $A_{[i,i]}$ only (*block Galerkin*), i.e.,

$$\tilde{A}^C = I^T A_u I \,, \tag{5.86}$$

with the block diagonal matrix:

$$A_u = \begin{pmatrix} A_{[1,1]} & 0 & \cdots & 0 \\ 0 & A_{[2,2]} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_{[M,M]} \end{pmatrix} .$$

The resulting coarse-grid matrix $\tilde{A}^C$ is block diagonal and thus sparser than the full Galerkin product. Hence, the computational cost on the coarser levels can be reduced, but the variational principle of Theorem C.6 is violated.

As pointed out, the unknown-AMG approach treats each physical unknown separately. Hence, the coarse-grids, transfer operators and coarse-grid operators can be adapted to the specific properties (e.g. anisotropies, coefficient jumps, ...) inside each physical property, as these are reflected within the diagonal matrix blocks $A_{[i;i]}$. Also, it is not required that the different physical unknowns are discretised on a common mesh. The downside is that information between different unknowns (the entries inside the off- diagonal block matrices $A_{[i;j]}$, $i \neq j$) is completely ignored. If these entries are large, the resulting AMG hierarchy may loose its efficiency as not all relevant information is reflected. The convergence statements analogous to those of Sec. C.4.4 are:

Lemma 5.3. *Let $A > 0$ and a VU mapping be given. If the C/F-splitting and interpolation I_{FC} are such that the τ-condition of MU-interpolation* (C.32)*,*

$$\|e_F - I_{FC} e_C\|^2_{0,F} \leq \tau_u \|e\|^2_{u,1} \tag{5.87}$$

is fulfilled with τ_u being independent of e and using the modified norm $\|v\|_{u,1} = (A_u v, v)_E$, then (C.31) *is satisfied with $\tau = \tau_u \rho(A^{-1} A_u)$ and with $K = I - I A_C^{-1} R A$ denoting the two-grid correction operator.*

Together with the smoothing property (C.22) the following result for a two-level cycle with one post-smoothing step is obtained.

Lemma 5.4. *Let $A > 0$ and a VU mapping be given and let S satisfy the smoothing property* (C.22)*. Furthermore, assume the C/F-splitting and interpolation to be such that condition* (5.87) *is fulfilled with some τ_u being independent of e. Then:*

$$\|SK\|_1 \leq \sqrt{1 - \frac{\sigma}{\tau}}\,,$$

is satisfied with $\tau = \tau_u \rho(A^{-1}A_u) \geq \sigma$.

The smoothing property (C.22) does not reflect the decomposition of A into the matrix blocks $A_{[i,j]}$. However, the AMG hierarchy is built solely using information from the diagonal blocks $A_{[i,i]}$. More precisely, implicit use is made of heuristics that rely on the algebraic smoothness of the error with respect to each $A_{[i,i]}$ separately. Hence, to obtain an efficient interplay between smoothing and coarse-grid correction, the smoothing property should be formulated accordingly.

Definition 5.5. *An operator S satisfies the unknown-smoothing property with respect to $A > 0$ and a given VU mapping, if a $\sigma_u > 0$ exists such that for all e:*

$$\|Se\|_1^2 \leq \|e\|_{1,u}^2 - \sigma_u \|e\|_{2,u}^2\,. \tag{5.88}$$

Here, the scalar product $(u, v)_{2,u} := u^T A_u^T D^{-1} A_u v$ and the associated norm $\|\cdot\|_{2,u}$ are defined with regard to A_u. Now, the two-grid correction operator with one post-smoothing step SK can be estimated in terms of $\|\cdot\|_{u,1}$ instead of $\|\cdot\|_1$.

Theorem 5.6. *([65], Theorem 3.9) Let $A > 0$ and S satisfy the smoothing property* (5.88)*. Furthermore, assume the C/F-splitting and interpolation to be such that the τ_u-condition of MU-interpolation* (5.87) *is fulfilled with τ_u being independent of e. Then:*

$$\|SK\|_{u,1} \leq \sqrt{\rho(A^{-1}A_u)\rho(A_u^{-1}A)}\sqrt{1 - \frac{\sigma_u}{\tilde{\tau}}}\,, \tag{5.89}$$

with $\tilde{\tau} = \tau_u \rho((A^{-1}A_u)^2)\rho(A_u^{-1}A)^2$.

The factor:

$$\rho_u := \rho(A^{-1}A_u)\rho(A_u^{-1}A)\,,$$

can be interpreted as an indication of the strength of unknown cross-couplings, i.e. how well the spectrum of A is captured by A_u. If ρ_u is large, the overall convergence may detoriate. Note, however, that the bound (5.89) is not sharp.

The essential conditions for the unknown-based approach to work are that, for each unknown, the submatrix of A reflecting the couplings of this unknown to itself is close to being an M-matrix and that smoothing results in an error which is smoothed separately for

each unknown. Advantages of this approach are that it can easily cope with anisotropies which are different between the different unknowns and that unknowns can be distributed arbitrarily across mesh points. This simple approach works quite efficiently for some important applications. However, it may become inefficient if the cross-function couplings are too strong.

5.4.4 Validation of the AMG approach

In what follows different configurations of the algebraic multigrid scheme are thoroughly validated using the standard lid-driven cavity problem, a test case that has been used extensively in the development of multigrid methods [119, 181, 259, 282]. The focus is on matrices formed by quadratic FE discretisation, which are certainly more complicated to solve than those arising from their linear element counterparts, and for this case the matrices resulting in the case of Stokes flow and those resulting from Newton linearisation of inertial flow are investigated separately. In the course of the analysis various influential factors such as the choice of least-squares weighting and Re are considered; furthermore the scaling performance of increasing numbers of variables is analysed. Note, that a symmetrised F-cycle is used, see Fig. 5.21d, explaining the higher runtimes compared to a W-cycle. The standard non-symmetric F-cycle which performs a further V-cycle only after each interpolation, returns run-times very close to those of the W-cycle.

(a) Lid-driven cavity; Stokes flow

Consider the 2D lid-driven cavity problem in the unit square, cf. Fig. 5.17, discretised using an unstructured, boundary-refined triangulation consisting of 27053 grid points and 13328 elements (for quadratic basis polynomials) resulting in a square system matrix with 108212 unknowns and of the order of 4.61 million entries, cf. the first column of Table 5.3. Note, that for all the tests carried out the least-squares weighting factors appearing in equations (5.12)-(5.21) were chosen to be $w = [1, 1, 5, 1]$ which is explained subsequently.

Initially, the algebraic smoothing properties of standard relaxation methods on the error of the above linear system are considered. While a simple Jacobi iteration proves to be totally inefficient and ILU-type smoothers are too costly, the GS method comes close to the desired and also expected behaviour. Table 5.2 shows the evolution of different error norms for i GS sweeps; the 2-norm decreases much faster than the 1-norm such that $\|e^{(i)}\|_{1,u} \ll \|e^{(i)}\|_{1,u}$ for increasing i as suggested in Sec. C.4.1(a). A comparison of the first with the third column shows that the unknown-smoothing property (5.88) is fulfilled with $\sigma_u \approx 0.92$. With increasing difference in the error norms the error smooths algebraically in the sense that residual reduction becomes inefficient as seen in the last column: while the first smoothing step reduces the residual significantly by a factor of 1/3 the following

sweep i	$\|e^{(i)}\|_{1,u}$	$\|e^{(i)}\|_{2,u}$	$\|e^{(i-1)}\|_{1,u} - \sigma_u \|e^{(i-1)}\|_{2,u}$	$\|r^{(i)} - Ae^{(i)}\|_E$
0	681.5375	519.4413	-	52.5562
1	203.1415	22.2745	203.6515	18.7161
2	155.9838	11.1301	182.6490	13.1397
3	130.3514	7.1448	145.7442	10.3753
4	113.8123	5.1305	123.7781	8.6982
5	102.0534	3.9132	109.0922	7.5414
6	93.1644	3.1018	98.4532	6.6826
7	86.1552	2.5312	90.3107	6.0187

Table 5.2: Evolution of different error norms for i GS sweeps.

steps have a progressively weaker impact until the residual tends to stagnate. Certainly only very few smoothing steps are effective while the remaining error components must be reduced by coarse-level correction.

For the coarse-grid correction three basic variants were tested; first of all the pure *scalar* AMG, that is disregarding the fact that a PDE depending on several physical unknowns is being dealt with; the second and third are the unknown-based PDE approach of Sec. 5.4.3 involving either the full Galerkin product (5.85) or the block Galerkin product (5.86), respectively. In the full Galerkin variant, at each level coarsening and interpolation are based on the block diagonal matrix only while the Galerkin product and also the smoothing procedure involve the full matrix. By this means also the coupling information *between the physical unknowns* is transferred to the coarsest level which is more expensive but might be beneficial for problems with a stronger unknown to unknown coupling. On the other hand, the block Galerkin variant is based solely on the block diagonal matrices while the coupling information only comes into play through the residual calculation at the finest level and by a possible acceleration treatment via, for example, the CG method.

The above three variants are used in combination with three different types of MG-cycle, the V, W and F-cycle, and are either CG-accelerated according to Alg. 6 or not. On each level between one and three pre-smoothing and post-smoothing steps are performed via a GS-relaxation method. At the coarsest level, i.e. when the number of unknowns falls below 2000, a direct solver is employed. The interplay between smoothing and coarse-grid correction for a case with fewer variables and full Galerkin AMG is shown in Fig. 5.22. Here the error in the cavity velocity in the x-direction is displayed at different levels before and after smoothing. Although the coarse-level variables are no longer connected to any kind of triangulation, for reasons of visualisation they have been connected by an artificial Delauny triangulation for levels 2 and 3. For convenience only, the boundary points are transported to the coarser levels, although they would have been sorted out into the F-variable set directly at the first level.

method	level 1	level 2	level 3	level 4	setup	CG+Vcyc
scalar					2.77 s	0.10 s
matrix size	108212	7478	1745	-		
number entries	4619322	422634	6661	-		
full Galerkin					1.39 s	0.13 s
matrix size	108212	13398	2310	1625		
number entries	4619322	1483123	97130	3031		
block Galerkin					0.89 s	0.05 s
matrix size	108212	13398	2310	1625		
number entries	1208892	373042	25512	1951		

Table 5.3: The relative complexity of the three AMG variants utilised subsequently; the setup time and time per cycle for each case is shown.

Attention is paid to the precise coarsening algorithm and the construction of interpolation operators which via the τ_u-factor from (5.87) mainly contribute to the convergence estimate of Theorem 5.6. A look at the matrix blocks (5.12)-(5.21) reveals that the diagonal blocks only depend on second order derivatives and will come close to M-matrices, see Definition C.3, with zero row-sums except for the rows belonging to boundary-adjacent nodes. So the conclusions of Sec. C.4.1(d), that the error evolves slowly along strong negative couplings, remains essentially valid, at least for couplings between nodes *of the same physical unknown* which is relevant for unknown-based AMG. Certainly the diagonal blocks are invertible with dominating positive diagonal entries but besides the majority of negative off-diagonal entries there will also be some comparatively small positive ones which are typical for higher order discretisations and arise for instance when using second order FE basis functions in contrast to first order ones; the latter becomes clear when analysing the couplings between the triangular shape functions (5.52)-(5.54).

When considering the origin of the positive entries there is no reason to assume an oscillatory error behaviour along the positive couplings so that it is reasonable to treat the positive and negative off-diagonal entries during the coarsening process in a completely symmetric way, as also suggested in [242]. Thus, all couplings are considered strong unless their *absolute* value undercuts a certain value which in the present implementation is for each row a fraction of $\varepsilon_{str} = 10^{-3}$ of the largest off-diagonal entry, cf. (C.25). If, however, the coarsening strategy is built upon the *whole* matrix, as in the case of scalar AMG, the off-diagonal matrix blocks come into play also. Then the M-matrix structure can no longer be assumed, also mixed derivatives are involved, and very large off-diagonal entries are present. Here, no simple arguments allow for conclusions on the relation between couplings and error behaviour and the best strategy is to simply treat all off-diagonal entries equally.

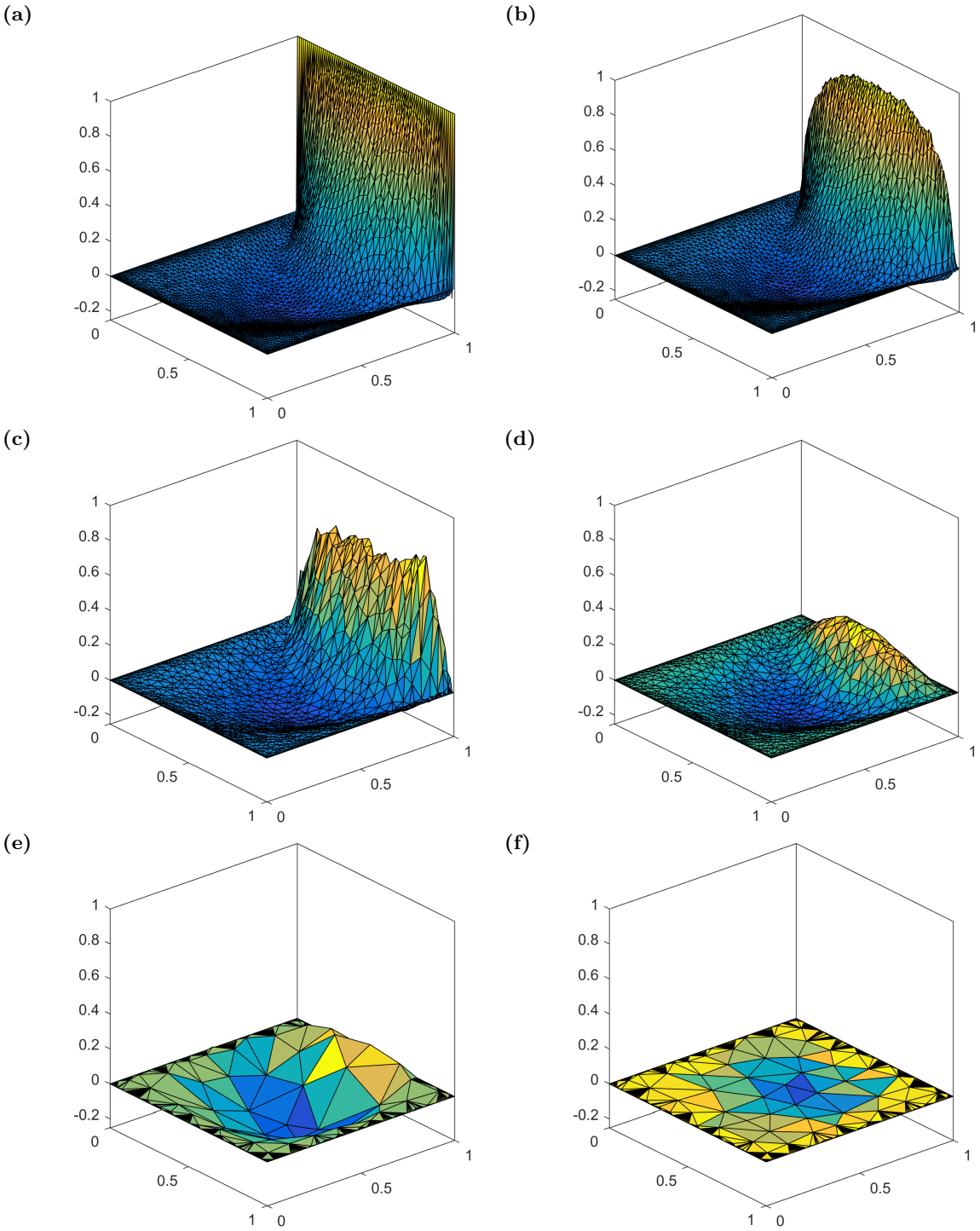


Figure 5.22: AMG error associated with u_x-velocity component during the restriction process; the AMG uses a full Galerkin approach with standard interpolation and at each level 20 smoothing steps are performed. Rows 1 - 3, from top to bottom, show the error at the first, second and third level of the AMG hierarchy (containing approximately 800, 400 and 100 grid points) demonstrating the error state before (left column) and after (right column) smoothing. Normally, boundary grid points are moved to the F-set at the first level; however, they have been transferred to the coarser levels here for reasons of visualisation.

case	AMG appr.	cycle	accel.	interp.	# sm.	# iter.	runtime [s]
C1	-	-	CG	-	-	5309	58.9
C2	full Galerkin	V	-	stand.	2	485	53.08
C3	full Galerkin	W	-	stand.	2	203	28.62
C4	full Galerkin	F	-	stand.	2	70	37.97
C5	full Galerkin	V	CG	stand.	2	43	6.77
C6	full Galerkin	W	CG	stand.	2	34	5.58
C7	full Galerkin	F	CG	stand.	2	17	10.81
C8	full Galerkin	V	CG	direct	2	652	127.17
C9	scalar	V	-	stand.	2	> 5000	> 500
C10	scalar	V	CG	stand.	2	768	77.11
C11	block Galerkin	V	CG	stand.	2	69	4.21
C12	block Galerkin	W	CG	stand.	2	59	3.84
C13	block Galerkin	F	CG	stand.	2	50	8.65
C14	block Galerkin	W	CG	stand.	1	74	3.59
C15	block Galerkin	W	CG	stand.	3	58	4.78

Table 5.4: AMG performance for the solution of the 2D lid-driven cavity problem using different solver ingredients. The labels "scalar", "full" and "block Galerkin" refer to the scalar AMG method (disregarding the fact that a PDE is being solved for), and the two versions of the unknown-based approach involving either the full Galerkin product (5.85) or the block Galerkin product (5.86). The abbreviations # sm. and # iter. correspond to the number of smoothing steps per level and the total number of iterations needed to reach an absolute residual smaller than $\epsilon = 10^{-8}$.

Generally, the error evolution *between* the physical unknowns will be non-smooth from a geometric point of view. Nevertheless, the variational principle of Theorem C.6 guarantees also in the case of scalar AMG, convergence of the method *regardless* of the precise form of coarsening and interpolation.

In all cases the standard coarsening procedure is complemented by either direct or standard interpolation, see Sec. C.4.3, which only needs minor modification in order to avoid negative interpolation weights in the case of positive off-diagonal entries; in principle, for all rows, the sign of positive off-diagonals can be assigned as negative in order to use the stated algorithms and then the diagonal entry modified accordingly in order to retain a zero row-sum. In the case of standard interpolation the interpolation operator is truncated by a factor of $\varepsilon_{tr} = 0.01$ according to the considerations of Sec. C.4.3(b) in order to avoid Galerkin operators that are too dense.

Table 5.3 shows the relative complexity of the three different AMG variants considered; in each case standard coarsening and standard interpolation is used. The coarsening

procedures and thus the sets of coarse-grid variables in the full and block Galerkin case are equivalent since they are based on the diagonal block structure only; however, the number of matrix entries differ depending on how the Galerkin product is constructed leading to significantly less entries in the block Galerkin case, a fact which is also reflected in the corresponding combined AMG setup and single cycle (involving one CG step and one AMG V-cycle) times. Here it is noted that the higher cycling time of the full against the block Galerkin variant is not only due to the more complex construction of the Galerkin product but also by the significantly higher cost of full relaxation sweeps versus block diagonal relaxation sweeps.

In all cases the ratio between the number of entries and the matrix size is strongly reduced if the initial system is compared to the system at the coarsest level: in the full Galerkin case this ratio is reduced from 42.7 to 1.87 and in the block Galerkin case from 11.17 to 1.20. That is to say, not only the matrix size reduces for each coarser level but also an increasingly sparser matrix structure is enforced; so that at the coarsest level the direct solver has to deal with an extremely sparse matrix only. The complexities of Table 5.3 demonstrate, for the standard coarsening algorithm of Sec. C.4.2(a), an absolutely sufficient coarsening speed with descending levels such that more aggressive strategies are unnecessary. This, amongst other possibilities, is due to the fact that quadratic FEs lead to a denser adjacency structure of the matrix so that for each C-variable more coupled neighbours can be moved into the F-set; however, in the case of linear elements a more aggressive strategy might prove useful.

Attention is now directed to the corresponding AMG performance results. Table 5.4 shows both the number of iterations and runtime required in order to reach an absolute residual smaller than $\epsilon = 10^{-8}$. Additionally Fig. 5.23 provides the associated convergence histories. Note that the underpinning Matlab code was not optimised and that the run times therefore do not reflect the full potential of the method; they are merely displayed to allow for a *relative* comparison between different AMG strategies.

A clear observable is that despite the favourable positive definiteness of the system matrix it is found that a pure conjugate gradient solver (C1) needs over 5000 iterations to satisfy the residual requirement; however, the iteration steps are computationally cheap resulting in a runtime of approximately 60 seconds. Although each of the AMG variants tested needs significantly fewer numbers of iterations than pure CG, a higher effort associated with the AMG cycles does not always pay off; at first glance two factors stand out. On the one hand, standard interpolation improves over direct interpolation (C8) to such an extent that the slight additional computational cost involved appears more than worthwhile. This result can essentially be explained in terms of the two-level convergence theory, see Lemma C.10, which requires the fine-level variables to be interpolated from

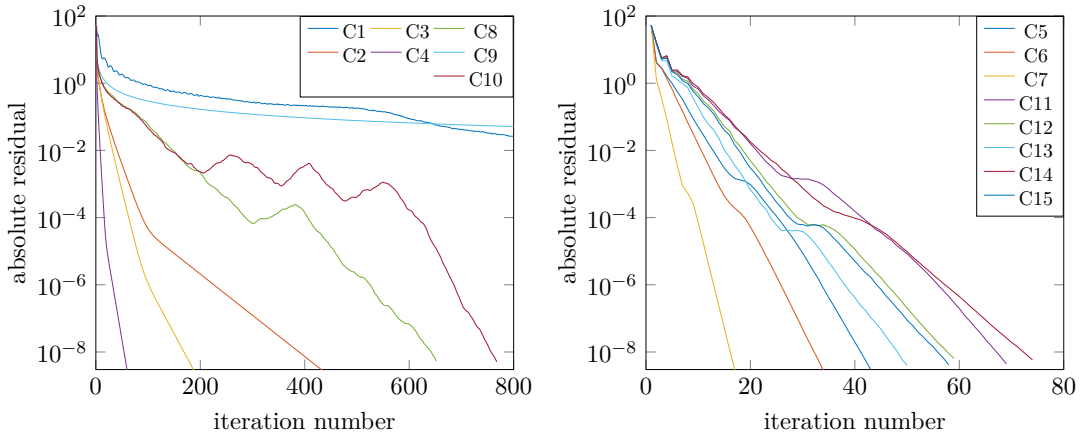


Figure 5.23: Absolute residual reduction of the AMG configurations of Table 5.4 for the 2D lid-driven cavity problem; the convergence history of the pure conjugate gradient solver (C1), of the non-accelerated AMG cycles (C2-C4) and of two particularly poor performing variants involving CG-accelerated AMG with direct interpolation (C8) instead of standard and CG-accelerated pure scalar AMG (C9) [left]; the same for the various variants of CG-accelerated AMG with different types of cycles and using either the full or the block Galerkin product, cf. Table 5.4 [right].

a fixed percentage of the surrounding neighbours; a criterion which is considerably better fulfilled by standard interpolation than by direct interpolation. The second obviously poor performing variant is scalar AMG (C9 and C10). Although the stand-alone multigrid solver converges smoothly as predicted by the variational principle of Theorem C.6 the convergence rate is similar to that of pure CG; it is improved, but insufficiently, by multigrid acceleration. It is observed that the wiggles appearing in the convergence curve of C10 are introduced merely by the behaviour of the acceleration method. For both variants the slow convergence can be explained by the fact that the system matrix as a whole does not exhibit the desired M-matrix property while the diagonal blocks do. Generally it seems also reasonable to expect the relaxation method, in this case a GS iteration, to have a strong smoothing effect on each physical unknown *seperately*, rather than a strong algebraic smoothing effect on the whole system. Indeed, a coarsening and interpolation strategy based only on the block diagonal information proves to be much more efficient.

The focus henceforth is directed at those AMG strategies that show promise of further improvement. In contrast to C9 and C10, the plain non-accelerated AMG cycles, C2-C4, in combination with standard interpolation exhibit an exemplary linear convergence except for the presence of a distinct kink occurring roughly at the mid-point of the iteration process, which is particularly noticeable in the V-cycle case, cf. Fig. 5.23. This indicates that the coarsening and interpolation processes, in principle, allow for an efficient transfer of relevant information between different levels; however, after successive reduction of the

"well captured" error components, the method is eventually left with error components which are represented by the AMG setup only sub-optimally leading to a slow-down of convergence. This effect can be decreased by using more elaborate cycling processes such as W- and F-cycles reducing the number of iteration steps from almost 500 to little more than 200 and even to 70, which although positively affecting the run times remain unsatisfactory.

C5-C7 confirm what is also known from the literature [242]: AMG acceleration or, in other words, the preconditioning of an iterative method by a multigrid scheme, nearly always pays off. While it is hardly possible or at least inefficient to optimise the interplay between smoothing and coarse-grid correction such that all relevant error components are well captured in the reduction process, the normally few remaining components, which belong to the remaining eigenvalue clusters of the multigrid preconditioned matrix, can be perfectly well eliminated by a standard Krylov solver such as CG. For example in the V-cycle case CG acceleration yields a reduction in the number of iterations and in the runtime by roughly a factor of 10. Also Fig. 5.23 demonstrates much more uniform convergence rates in the accelerated cases where the pronounced kinks are reduced to small disturbances after which the typical fast convergence speed is restored.

Comparison between the three full Galerkin variants C5-C7 and the block Galerkin variants C11-C13 reveals that, at least for the matrices resulting in the Stokes flow case, the time saving of the latter variants slightly exceed the extra cost of more cycles, even though there are 50 compared to 17 cycles in the F-cycle case for example (C13 versus C7). A glance at the convergence histories shows quantitative differences between the full and block Galerkin schemes only, no qualitative differences. In general the performance of V- and W-cycles is very similar, the latter slightly outperforming the former; on the other hand, F-cycles always lead to by far the lowest number of iterations but their use does not seem to altogether pay off. Finally, application of a single pre- and post-smoothing step at each level seems to achieve sufficient smoothing for the further coarsening such that any additional relaxation sweeps are not cost effective. The best AMG variant (C14) proves to be about 15 times faster than a pure CG iteration.

(b) Lid-driven cavity: inertial flow

The AMG's performance is investigated for linear systems arising in the case of Newton-linearisation of inertial flow: starting with the Stokes flow solution as an initial guess, the performance data provided in Table 5.6 are for the first Newton iteration which is assumed to be representative for all subsequent iterations. Unless stated otherwise $\mathrm{Re} = 200$; use is made of an unstructured, boundary-refined triangular grid having the same number of grid points, elements and unknowns as for the Stokes flow case, but due to the additional matrix entries stemming from Newton linearisation the total number of non-zero entries,

method	level 1	level 2	level 3	level 4	setup	CG+Vcyc
full Galerkin					1.39 s	0.1 s
matrix size	108212	13397	2315	1625		
number entries	4819994	1484313	98091	3031		
block Galerkin					0.89 s	0.04 s
matrix size	108212	13397	2315	1625		
number entries	1209214	372933	25767	1951		

Table 5.5: The relative complexity of two AMG variants for matrices resulting from Newton-linearised systems for inertial flow. The setup involves using standard coarsening and standard interpolation.

4.82 million, is approximately 4% higher than in the Stokes case. The complexities of the AMG setup for the full and block Galerkin cases are summarised in Table 5.5 and are seen to be close to those of Table 5.3. The coarsening procedures and thus the sets of coarse-grid variables in the full and block Galerkin case are equivalent since they are based on the diagonal block structure only; however, the number of matrix entries differs depending on how the Galerkin product is constructed leading to significantly less entries in the block Galerkin case. This is also reflected in the corresponding run times for the AMG setup and for a single cycle involving one CG step and one AMG V-cycle. Due to its inefficiency, scalar AMG is not considered a serious option and is therefore omitted from future comparisons.

For reasons of comparability the values in Table 5.6 were obtained under the same conditions as those in Table 5.4. Again, in all cases standard coarsening and standard interpolation was applied and at the coarsest level, i.e. when the number of unknowns becomes equal to or less than 2000, a direct solver is employed. The solver iterates until the absolute residual reaches a cut-off value less than $\epsilon = 10^{-8}$. The cases investigated in Table 5.6 correspond to those in Table 5.4, except that C8-C10 are excluded due to inefficiency while C14* and C15* are carried out with V instead of W-cycles. The least squares-weighting is applied in a different way: several benchmark tests for higher Reynolds numbers revealed that the new terms entering the matrix due to linearisation - which may even be of order Re^2, cf. equations (5.25)-(5.30) - lead to a disproportionate least-squares weighting between the equations of momentum conservation and those of mass conservation, thus necessitating change of the weighting factors. In order to avoid AMG convergence loss, a heuristic weighting of the continuity equation by a factor proportional to $\sqrt{\mathrm{Re}}$ proves adequate and subsequently the rule $w = [1\,1\,\max\{5, \sqrt{\mathrm{Re}}\}\,1]$ is adhered to, see Fig. 5.25.

An initial comparison between the linear and the-non-linear case, Tables 5.4 and 5.6, reveals that in the latter case the iteration numbers and run-times increase significantly.

While for the pure CG method this is true by a factor of about 3, the stand-alone multigrid variants perform decidedly worse. Although the shapes of the corresponding convergence curves in Fig. 5.24(left) are qualitatively similar to those in Fig. 5.23(left), the iteration numbers and run-times for C2-C4 are increased by factors of up to 10. Comparison between C1 and C2-C4 demonstrates the strong deterioration in the multigrid efficiency for the non-linear problem and that the use of acceleration is certainly necessary. Indeed, C5-C7 and C11-C13 show that CG-acceleration improves the performance dramatically pushing the cycle number ratio between the non-linear and linear examples below 3 and in the best case even below 2; nevertheless the corresponding convergence curves in Fig. 5.24(right) all exhibit a number of small plateaus which seem to stretch with growing Reynolds number, cf. also Fig. 5.25.

In contrast to the linear case the block Galerkin method does not dominate the full Galerkin variant but the result is inconclusive. Here a trend can be identified which continues for increasing Reynolds number: the cycle number ratio between the block and full Galerkin method does not remain constant but increases in favour of the full Galerkin method, an effect which is firstly noticeable for the more complex F- and W-cycles but will later also be apparent for the V-cycle. However, in the present case with Re = 200 the block Galerkin method is still the best but now, in contrast to the Stokes case, in combination with the cheapest variant using V-cycles. The best performance (C14*) achieved with one pre- and one post-smoothing step per level, is more than 20 times faster than pure CG and just 1.9 times slower than the best performer in Table 5.4.

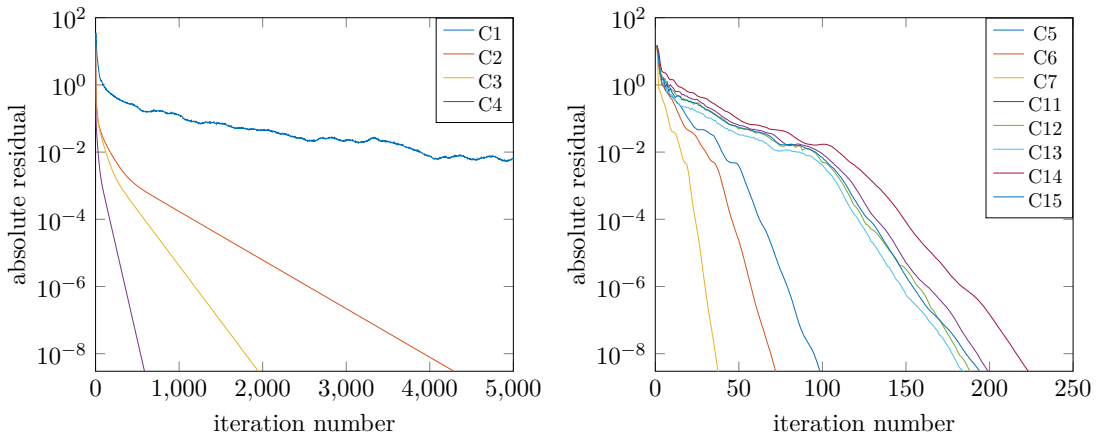


Figure 5.24: Absolute residual reduction of the AMG configurations of Table 5.6 for the non-linear 2D lid-driven cavity problem with Re = 200; the convergence history of the pure conjugate gradient solver (C1) and of the non-accelerated AMG cycles (C2-C4) (left); the same for the various variants of CG-accelerated AMG with different types of cycles and using either the full or the block Galerkin product (right).

case	AMG appr.	cycle	accel.	interp.	# sm.	# iter.	runtime [s]
C1	-	-	CG	-	-	13877	143.85
C2	full Galerkin	V	-	stand.	2	4771	471.99
C3	full Galerkin	W	-	stand.	2	2120	251.03
C4	full Galerkin	F	-	stand.	2	676	315.02
C5	full Galerkin	V	CG	stand.	2	100	10.84
C6	full Galerkin	W	CG	stand.	2	73	9.75
C7	full Galerkin	F	CG	stand.	2	38	18.98
C11	block Galerkin	V	CG	stand.	2	203	8.22
C12	block Galerkin	W	CG	stand.	2	192	9.46
C13	block Galerkin	F	CG	stand.	2	188	29.45
C14*	block Galerkin	V	CG	stand.	1	227	6.81
C15*	block Galerkin	V	CG	stand.	3	197	10.39

Table 5.6: AMG performance for the solution of the non-linear 2D lid-driven cavity problem, Re = 200: starting with the Stokes flow solution as an initial guess the performance is measured following the first subsequent Newton iteration which is assumed to be representative of subsequent iterations. In contrast to cases C14 and C15 in Table 5.4 the above cases C14* and C15* are carried out with V instead of W-cycles.

Fig. 5.25 shows the AMG residual history of C5 for different Reynolds numbers; based on the heuristic least-squares weighting $w = [1,\ 1,\ \max\{5, \sqrt{\mathrm{Re}}\},\ 1]$ (left) and the standard weighting $w = [1,\ 1,\ 1,\ 1]$ (right). It is clear that the effort required to solve the linear systems using Newton-iteration grows continuously with Re and that the type of least-squares weighting strongly influences the effective dependency: for standard weighting computational effort versus Re does not scale proportionally, unlike the heuristic weighting which does; in the latter case the computational cost is reduced by a factor of up to three and the convergence curves themselves are comparatively smoother. A broader investigation of the influence of weighting on AMG performance reveals that any choice, other than the over-weighting of only the continuity equation, downgrades the performance; this is particularly the case if the third and fourth equation in the system (4.12) are weighted equally, as in Sec. 5.3.1(c), when investigating the problem of flow past a backward-facing step.

Fig. 5.26 (left) shows, for the AMG configuration C5 of Table 5.6, the number of iteration steps necessary to solve a linearised equation using the Newton method versus the least-squares weighting factor w_c of the continuity equation. Here, only the weighting factor for the continuity equation is varied while the other weighting factors are kept constant, that is the overall weighting is $w = [1,\ 1,\ w_c,\ 1]$. Obviously the curves exhibit

distinct minima which are significantly below the equal weighting case $w_c = 1$; these minima shift to the right-hand side with increasing Re, suggesting as earlier a choice for the weighting factor as proportional to $\sqrt{\text{Re}}$. It should be borne in mind that in order to solve a very high Reynolds number problem, it might be more efficient to utilise a stepping algorithm with moderate Re steps or to increase Re during the Newton-iteration; in such cases only moderate weighting would be necessary.

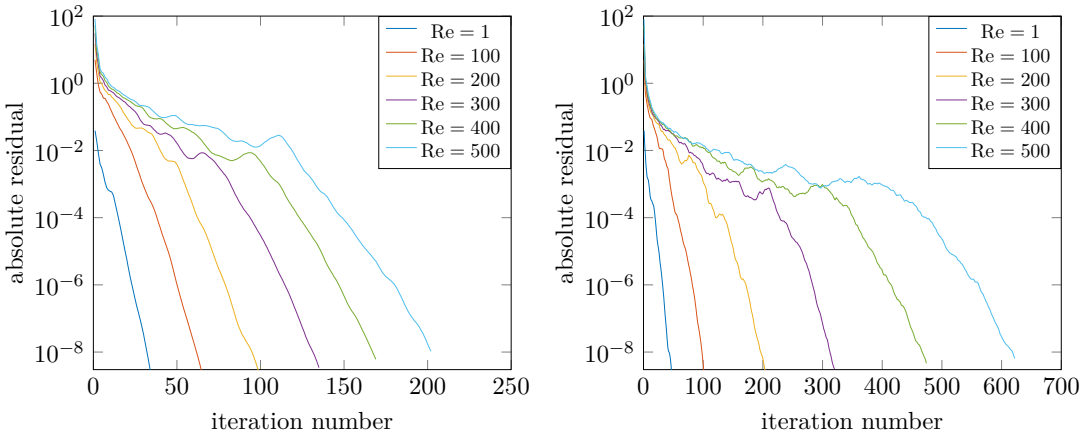


Figure 5.25: Absolute AMG residual reduction for C5, Table 5.6, and for different Reynolds numbers based on: heuristic least-squares weighting $w = [1, 1, \max\{5, \sqrt{\text{Re}}\}, 1]$ (left); standard weighting $w = [1, 1, 1, 1]$ (right).

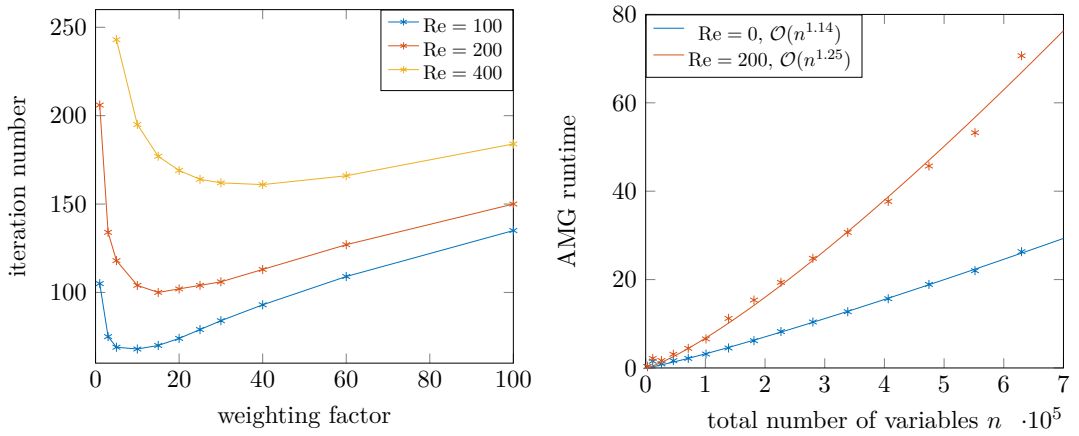


Figure 5.26: Number of iterations necessary to solve C5 depending on the least-squares weighting factor for the continuity equation w_c and an overall weighting of $w = [1, 1, w_c, 1]$, when Re = 100, 200, 400 (left). AMG run-time versus total number of variables n [problem size] revealing the scaling performance of the AGM; the runtime in seconds required to solve the Stokes problem (blue line) and to obtain a solution of the first subsequent Newton-iteration (red line) for Re = 200 taking the Stokes matrix as initial guess (right).

The last and one of the most important aspects addressed in the present analysis is the question of how the performance of the AMG scales with increasing problem sizes. The main reason for using a multigrid approach is the possibility of achieving *linear* scaling as reported in the literature. Fig. 5.26 (right) investigates the runtime of the AMG approach when solving the Stokes problem followed by the first Newton-iteration of the non-linear problem with Re = 200. In both cases the best performing variants from Tables 5.4 and 5.6 are used, that is C14 in the linear and C14* in the non-linear case. Run-time is plotted against the matrix size in which the latter varies from a few thousand to over 600 thousand with increasing problem size. In the non-linear case the largest matrix involves approximately 630,000 variables, over 28.5 million entries and a condition number of 1.2×10^9. The measured run-times were fitted to a function of the form $f(n) = c \cdot n^k$ via a non-linear least-squares algorithm revealing what is also clear to the naked eye: the scaling law is slightly over-linear with exponents of 1.14 and 1.25 in the Stokes and the non-linear case, respectively. These exponents might be further improved by optimising individual multigrid components, but for the purposes of the present work are acceptable.

Although the above analysis is restricted to a specific flow problem, the 2D lid-driven cavity, it can be argued most of the results will transfer, in one way or another, to arbitrary viscous and inertial flow problems. The suggested multigrid configuration using a block Galerkin approach, GS smoothing, standard coarsening and interpolation, V- or W-cycles and CG-acceleration, in combination with a suitable least-squares weighting, forms a satisfactory basis for defining an efficient black box method for the solution of linear systems arising from a least-squares FE-discretisation. The caveat being that in the context of overall implementation, scope for improvement remains.

6 Engineering applications

In this chapter a series of problems involving lubrication and film-like flows are investigated which appear and play a key role in a number of industrial applications. The problems are solved using the new FE method and solution strategies described in Chapter 5 complemented by more customised series solutions based on complex-valued description when applicable. The whole is centred on the first integral methodology in order to demonstrate its versatility, efficiency and convenience in addressing a range of flow types.

The industrial applications are chosen in an order of increasing complexity, starting with hydrodynamic lubrication between corrugated surfaces moving relative to each other, an archetypal form of friction which is modelled in a 2D steady Couette-like setting, see Sec. 6.1. While this problem deals with periodic and wall boundary conditions only, the convenient form of the dynamic boundary condition comes into play when considering steady coating flows over planar and curvilinear surfaces as explored in Secs. 6.2 and 6.3; the latter explicitly demonstrating that the advantages related to 2D steady film flow extend to arbitrary axis-symmetric flows. For the solution of periodic film flows a spectral solution method is provided which lend itself to predicting the stability of film flows in particular.

One aspect of lubrication applications with corrugated surfaces is the process of material exchange between the fluid inside the valleys of the corrugations and the bulk flow which, in some applications, it is desirable to control, for instance to achieve a minimum mixing of the lubricant fluid, see Sec. 6.4. This problem is intrinsically unsteady requiring an approach different to the above; in this context a generalisation of the complex-valued Goursat formulation and unsteady FE formulation, together with corresponding computational results, are presented.

6.1 Friction reduction between lubricated surfaces

6.1.1 Problem description and solution approach

Modern combustion engines have to comply with ever increasing challenges, whether with respect to performance attributes and efficiency, or the emission of pollutants. One of numerous approaches to increase their efficiency is the optimisation of the tribological

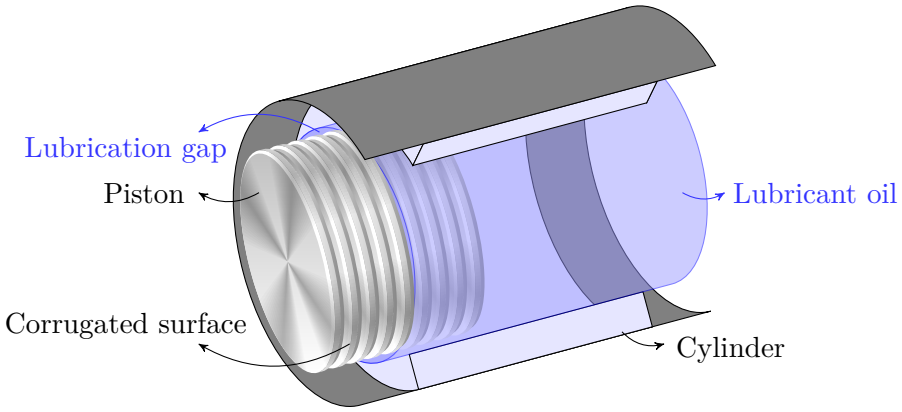


Figure 6.1: Schematic of a cylinder-piston system: the cylinder wall (light gray) has a smooth surface while the piston skirt (metallic) is corrugated with the objective of reducing the hydrodynamic friction in the lubricated (blue) gap as the cylinder and piston move relative to each other.

processes involved in lubricated contacts. The friction generated between moving surfaces is complex since both dry and hydrodynamic friction can occur depending on the conditions, normally a mixture of both ensues. In the case of hydrodynamic friction its value in a lubricated contact, as a function of the relative speed of the moving parts, is usually described in the form of a Stribeck curve [241].

If one takes the cylinder-piston system as an example, see the schematic of Fig. 6.1, experience shows that the friction generated is strongly influenced by the precise shape of the surfaces involved with experimental and numerical investigations being strongly focused on optimising the shape of the piston ring [249, 280]. With increasing optimisation components such as the piston skirt have come into focus. Typical methods to reduce friction is the reduction of microscopic surface roughness by grinding processes and/or periodically profiling the surface on a more macroscopic level; the aim of the latter being to decrease the effective area generating the hydrodynamic friction in contrast to that generated by a flat surface; cf. the piston profile in Fig. 6.1 shown up close as a cross-section of the lubricated gap in Fig. 6.2. Profiles such as these can promote the generation of eddies in the corrugation valleys which in effect act as a hydrodynamic ball bearing as is known from work relating to the shark skin effect [83]. Increasing the clearance (separation of the two surfaces) can also reduce friction but in many practical problems is not a realistic option. Accordingly, in the present study profiled moving surfaces containing a suitable periodically repeating structure is considered, as these can be realised practically.

Diverse geometric parameters are involved, other than the specific type of corrugation contour; in particular the amplitude, i.e. corrugation depth, and wave length, i.e. number of corrugations per unit length. Depending on these parameters and considering relevant

manufacturing constraints a friction coefficient can be measured and minimised to find a suitable corrugation structure. To obtain a complete picture the assessment of surface profiling should be based on a systems behaviour in the hydrodynamic and mixed regime; however, it is useful to decouple the two different effects and, as a first step, to concentrate on investigating the pure hydrodynamic regime as is the case here. For this purpose a simple two-dimensional model for shear flow over a corrugated surface is considered using fluid flow models of increasing complexity; these include the lubrication equations, the Stokes and the NS equations. Thus also the range of validity for the first two methods is clarified. Following on from the work of [217, 221] who demonstrated that, based on the objective of reducing the wall shear stress, a periodically occurring ridge-like structure is more suitable than a corresponding sinusoidal variation, attention is focussed on the former which have received less attention.

6.1.2 Mathematical modelling

A realistic piston movement, see Fig. 6.1, is a time-dependent process and thus relevant quantities affecting friction, depending on the velocity gradients, will be time-dependent too. Even if dynamic temperature effects and variation away from parallel alignment are neglected, the piston velocity U will cover a wide range of values. Nevertheless, considering a small periodic length λ of the piston skirt profile in comparison to the piston length it is reasonable to assume the velocity field in the lubricating gap to be close to that of a set of steady states of the periodic shear flow problem which might depend on a range of Reynolds numbers defined by the approximate minimum and maximum values for those constants (characteristic length and velocity, density, viscosity) occurring in a realistic time-dependent process.

In order to achieve a rough guide and impression of how surface profiling might influence hydrodynamic friction, a two-dimensional shear flow model amenable to simple analytic methods, is explored. Consider a cross section of the lubrication gap, Fig. 6.2a, with a periodically contoured lower surface and a parallel-aligned flat upper surface moving with a relative velocity of U; the minimum and average gap height are denoted by H_0 and H, the periodic length of the bottom contour by λ.

Unless the lubrication approximation is employed, which allows for an analytic solution as shown below, solutions of the Stokes and full NS equations are obtained by the first integral formulation (4.12) and the least-squares FE technique provided in Ch. 5. Fig. 6.2 shows the problem setup, with physical quantities on the left-hand side (a), and non-dimensioned ones on the right-hand side (b), for which the corresponding scalings are given in Table 6.1. The physical coordinates $\tilde{x}$ and $\tilde{z}$, together with H_0, H and λ, are scaled by the inverse wave number of the bottom contour, i.e. $\lambda/(2\pi)$, and the velocities $\tilde{u}_1$, $\tilde{u}_2$ are

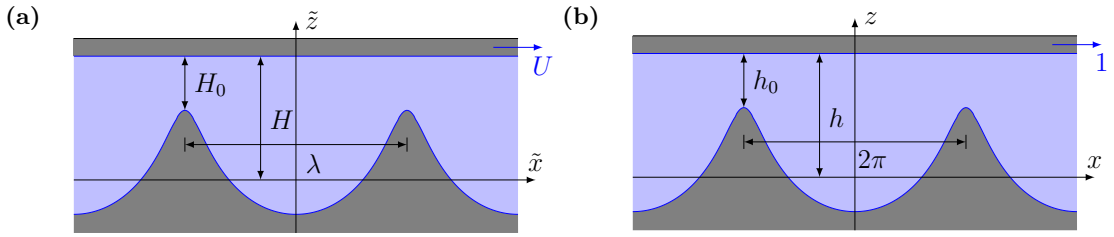


Figure 6.2: Shear flow geometry over a corrugated bottom contour showing (a) the physical and (b) non-dimensioned quantities.

physical quantities	dimensionless quantities	scaling factor
$\tilde{x}$, $\tilde{z}$	x, z	$\frac{\lambda}{2\pi}$
$\tilde{u}_1$, $\tilde{u}_2$	u_1, u_2	U
$\tilde{\phi}_1$, $\tilde{\phi}_2$	ϕ_1, ϕ_2	ηU
$\tilde{p}$	p	$2\pi\eta\frac{U}{\lambda}$
$\tilde{\psi}$	ψ	$\frac{U\lambda}{2\pi}$

Table 6.1: Scaling of the fundamental quantities: coordinate unit lengths, velocities, auxiliary potential fields, pressure and the streamfunction.

scaled by the lid velocity U, leading to the corresponding scalings of the potential fields ϕ_1, ϕ_2, the pressure p and the streamfunction ψ.

Accordingly, the non-dimensioned boundary value problem, using the operator notation from Sec. 4.2 with $u = (u_1, u_2, \phi_1, \phi_2)$, is:

$$\mathcal{L}u + \mathcal{N}[\mathrm{Re}, u]u = 0\,, \qquad \text{in the domain}\,, \tag{6.1a}$$

$$u_1(x, b(x)) = u_2(x, b(x)) = 0\,, \qquad \text{along the lower surface,} \tag{6.1b}$$

$$u_1(x, h) = 1\,, \quad u_2(x, h) = 0\,, \qquad \text{along the upper surface,} \tag{6.1c}$$

with the problem domain confined to the interval $x \in [-\pi, \pi]$ and $z = b(x)$ denoting the 2π-periodic lower surface contour; the problem domain is assumed periodic on the left, $x = -\pi$, and right-hand side, $x = \pi$, which is assured by identification of the respective grid points. Note that in contrast to the velocities, which are assumed 2π-periodic, the potential fields ϕ_1, ϕ_2 can exhibit an aperiodic part (being linear in x) as mentioned in Sec. 5.1.2.

The Reynolds number based on the above scaling is $\mathrm{Re} := \varrho\lambda U/(2\pi\eta)$. Considering standard piston velocities, typical viscosities for the occurring temperature range and manufacturing restrictions, the following characteristic parameter range is assumed [100]:

$$\varrho = 850\,\mathrm{kg/m^3}\,, \qquad 0.5\,\mathrm{mm} \le \lambda \le 1\,\mathrm{mm}\,,$$
$$20\,\mathrm{mPa\,s} \le \eta \le 60\,\mathrm{mPa\,s}\,, \qquad 0\,\mathrm{m/s} \le U \le 15\,\mathrm{m/s}\,,$$

leading to Reynolds numbers in the range of $0 \le \mathrm{Re} \le 100$.

Central to the present study is the definition of the friction coefficient depending on the non-dimensionalised quantities. For this purpose, note that the resistance force for the upper surface is given by the integral:

$$F = \iint_A \eta \left[\frac{\partial \tilde{u}_1}{\partial \tilde{z}} - \frac{\partial \tilde{u}_2}{\partial \tilde{x}} \right] \mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y} = \eta B \int_0^L \frac{\partial \tilde{u}_1}{\partial \tilde{z}}\,\mathrm{d}\tilde{x} = \eta B \frac{L}{\lambda} \int_{-\lambda/2}^{\lambda/2} \frac{\partial \tilde{u}_1}{\partial \tilde{z}}\,\mathrm{d}\tilde{x}$$
$$= \eta A \frac{2\pi}{\lambda} U \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial u_1}{\partial z}\,\mathrm{d}x = \eta \frac{A}{H_0} \left[\frac{h_0}{2\pi} \int_{-\pi}^{\pi} \frac{\partial u_1}{\partial z}\,\mathrm{d}x \right] U\,,$$

in which a comparison to pure Couette flow between two flat surfaces ($F = \eta A U / H_0$) identifies the quantity:

$$\kappa := \frac{h_0}{2\pi} \int_{-\pi}^{\pi} \frac{\partial u_1}{\partial z}\,\mathrm{d}x = \frac{h_0}{2\pi} \int_{-\pi}^{\pi} 2\frac{\partial \phi_2}{\partial x}\,\mathrm{d}x = \frac{h_0}{\pi} \left[\phi_2(\pi, h) - \phi_2(-\pi, h) \right]\,, \tag{6.2}$$

as the scalar factor describing the change of resistance in the presence of a curved surface contour in comparison to that of a flat surface; subsequently κ is referred to as the friction coefficient or resistance factor. The last equality in the above equation is obtained by the definition of the potential fields, see for example (5.39), (5.40), and demonstrates that κ can be obtained in a relatively simple and convenient way. In contrast to a formulation in terms of primitive variables, no derivatives have to be calculated from the FE solution and no integral needs be calculated; just the potential field ϕ_2 has to be evaluated at the two end points of the upper boundary revealing that κ only depends on the non-periodic part of ϕ_2.

6.1.3 Lubrication solution

Lubrication theory opens up the possibility of an analytical approximation. The starting point is the Reynolds equation which, for the present geometry, takes the form [217, 241]:

$$\frac{\mathrm{d}}{\mathrm{d}x} \left\{ \left[h - b(x) \right]^3 \frac{\mathrm{d}p}{\mathrm{d}x} \right\} = 6 \frac{\mathrm{d}}{\mathrm{d}x} \left[h - b(x) \right]\,, \tag{6.3}$$

which following integration with respect to x directly becomes:

$$\left[h - b(x) \right]^3 \frac{\mathrm{d}p}{\mathrm{d}x} = 6 \left[h - b(x) \right] - C\,,$$

with integration constant C. Solving for the pressure gradient gives:

$$\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{6}{[h-b(x)]^2} - \frac{C}{[h-b(x)]^3}\,, \tag{6.4}$$

which can be integrated over one period and, considering the periodicity of the pressure field, yields the identity:

$$0 = p(\pi) - p(-\pi) = \int_{-\pi}^{\pi} \frac{\mathrm{d}p}{\mathrm{d}x}\mathrm{d}x = 6\int_{-\pi}^{+\pi} \frac{\mathrm{d}x}{[h-b(x)]^2} - C\int_{-\pi}^{+\pi} \frac{\mathrm{d}x}{[h-b(x)]^3}\,, \tag{6.5}$$

with integration constant:

$$C = 6h\frac{I_2}{I_3} \quad \text{and} \quad I_n := \frac{1}{2\pi}\int_{-\pi}^{+\pi} \frac{\mathrm{d}x}{[1-b^*(x)]^n}\,, \quad b^*(x) := \frac{b(x)}{h}\,.$$

Following Scholle [217] an explicit form of the x-component of velocity is given by:

$$\begin{aligned} u_1 &= \frac{z-b(x)}{h-b(x)} - \frac{\mathrm{d}p}{\mathrm{d}x}\frac{[h-b(x)]^2}{2}\left[1-\frac{z-b(x)}{h-b(x)}\right]\frac{z-b(x)}{h-b(x)}\,, \\ &= Z - 3\left[1-\frac{I_2}{I_3\,[1-b^*(x)]}\right][1-Z]\,Z\,, \end{aligned} \tag{6.6}$$

with $Z := \frac{z-b(x)}{h-b(x)}$; which can be further integrated with respect to z to obtain the stream-function:

$$\begin{aligned} \psi &= \int u_1\,\mathrm{d}z = [h-b(x)]\int u_1\,\mathrm{d}Z \\ &= \frac{h-b(x)}{2}Z^2 - 3\left[h-b(x)-\frac{hI_2}{I_3}\right]\left[\frac{Z^2}{2}-\frac{Z^3}{3}\right] \\ &= h\left[\frac{3I_2}{2I_3}-1+b^*(x)\right]Z^2 + h\left[1-\frac{I_2}{I_3}-b^*(x)\right]Z^3\,. \end{aligned}$$

Obviously ψ vanishes along the contoured surface $Z = 0$ while taking a constant value along the upper flat surface $Z = 1$.

Accordingly, the non-dimensioned shear stress is given by:

$$\tau \approx \frac{\partial u_1}{\partial z} = \frac{1-3\left[1-\frac{I_2}{I_3[1-b^*(x)]}\right][1-2Z]}{h\,[1-b^*(x)]}\,,$$

in which case the shear stress along $Z = 0$ takes the form:

$$\tau_w \approx \frac{3I_2}{I_3 h\,[1-b^*(x)]^2} - \frac{2}{h\,[1-b^*(x)]}\,,$$

and along $Z = 1$ the form:

$$\tau_p \approx \frac{4}{h\,[1 - b^*(x)]} - \frac{3I_2}{I_3 h\,[1 - b^*(x)]^2}\,.$$

The corresponding resistance factor κ is given by:

$$\kappa \approx \frac{h_0}{2\pi} \int_{-\pi}^{\pi} \left[\frac{4}{h\,[1 - b^*(x)]} - \frac{3I_2}{I_3 h\,[1 - b^*(x)]^2} \right] \mathrm{d}x = \frac{H_0}{H} \left[4I_1 - \frac{3I_2^2}{I_3} \right], \tag{6.7}$$

depending on easily computable functionals of the contoured surface.

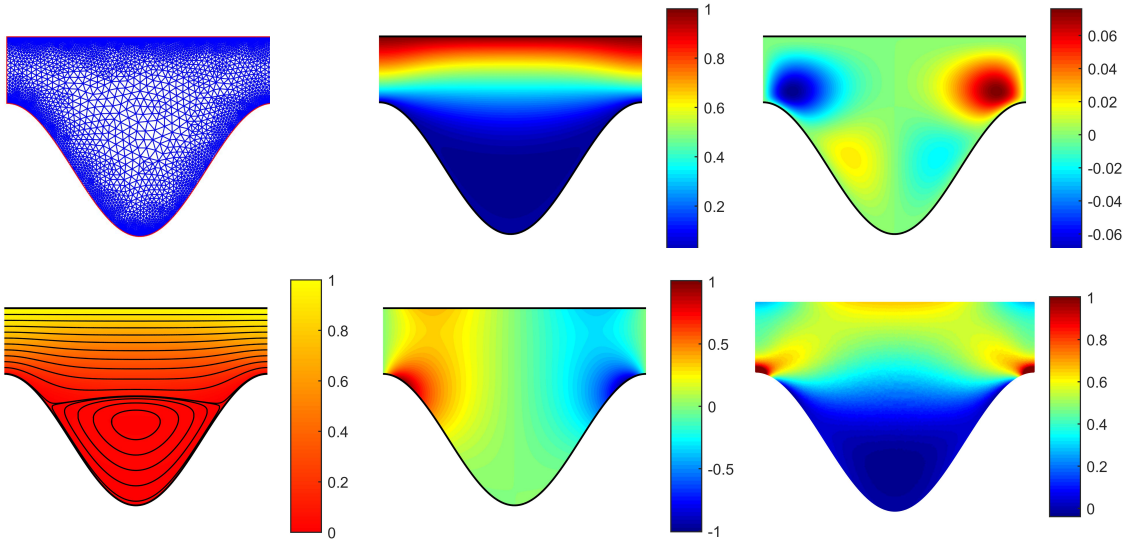


Figure 6.3: Stokes flow results for a cosine-like surface contour with $h_0 = a = \pi/2$. Top row (from left to right): unstructured FE-triangulation, x-velocity and z-velocity; bottom row (left to right): absolute velocity plus streamlines, pressure and scaled shear stress. The flow is from left to right.

6.1.4 A parametric study

The numerical results shown subsequently, for Stokes and NS flow, were obtained by the LSFEM of Ch. 5 using quadratic basis functions and an unstructured, boundary-refined triangular grid for all test and solution spaces and a least-squares weighting of $w = [1,\, 1,\, 10^4,\, 10^4]$; the boundary triangles are allowed to be curved in accordance with the isoparametric concept and the number of grid points involved ranged from 30,000 to 50,000. A typical Stokes solution using the surface contour $b(x) = -a\cos(x)$ is depicted in Fig. 6.3 which shows the corresponding unstructured FE triangulation, the x and z velocity components, the resulting streamline pattern, the pressure and the shear stress.

The latter two have been constructed from the velocities and potentials according to $p = \partial_k \phi_k - u_k u_k/2$, see equation (1.42), and $\tau = \partial u_1/\partial z$. Note, that the periodic boundary conditions require special treatment as described in Sec. 5.1.2.

(a) Wave number study

The impact of varying the wave number on hydrodynamic friction is now investigated. Taking, in non-dimensioned form, shear flow over a cosine-like contour with fixed constants a, h and h_0, the number of ridges can be increased by considering a sequence of contours $b_K(x) = -a\cos(Kx - \gamma\pi)$, $K \in \mathbb{N}_+$ on the interval $[-\pi, \pi]$ with $\gamma := 1$ for even K and $\gamma := 0$ otherwise; K then specifies the number of ridges per unit length 2π. However, in order to generate a continuous curve it is easier to consider the equivalent problem on the interval $[-\pi, \pi]$ with contours $b_K(x) = -aK\cos(x)$ and clearance $h_{0,K} = h_0 K$ with $K \in \mathbb{R}_+$. The associated friction coefficient κ is obtained using either formula (6.2) or (6.7) and, for three selected sequences of a and h_0, is displayed in Fig. 6.4. As expected, only in the limit case $K \to 0$ are the lubrication and Stokes results in agreement, lubrication theory only being valid if λ is the dominant length scale or in non-dimensioned quantities $a \ll 2\pi$. Obviously, lubrication theory results in constant friction for all K and is thus inappropriate for studying wave number effects. Alternatively, the Stokes equations qualitatively describe the entire transition from a near lubrication state ($K \to 0$) to a Couette-like state ($K \to \infty$) for arbitrarily dense ridge structures.

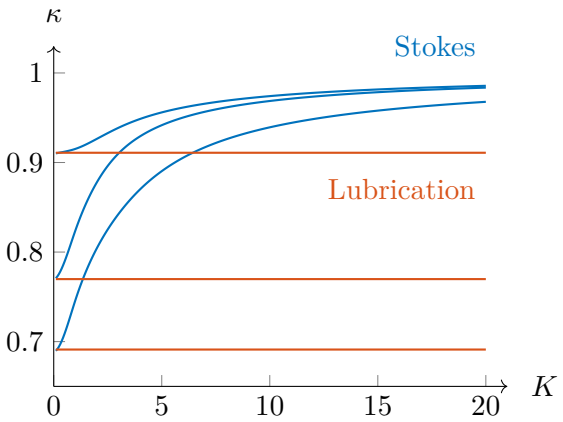


Figure 6.4: Plots of friction coefficient κ versus wave number K obtained with: the lubrication approximation (red, constant); the Stokes equations (blue). The pairs of curves shown are for: $a = K\pi/16$, $h_0 = K\pi/2$; $a = h_0 = K\pi/2$; $a = K\pi/2$, $h_0 = K\pi/4$.

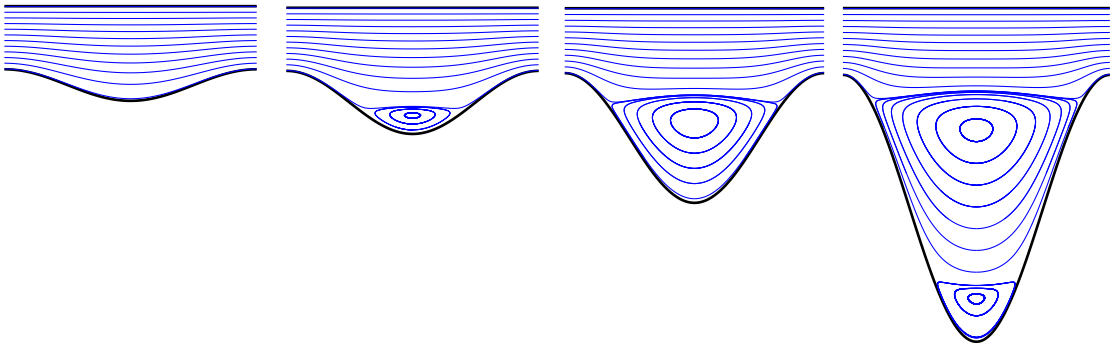


Figure 6.5: Vortex structure dependence on the amplitude of the cosine-like surface contour under Stokes conditions. Streamlines are depicted (from left to right) for fixed $h_0 = \pi/2$ and $a = \pi/8$, $\pi/4$, $\pi/2$ and π. Flow is from left to right.

(b) Amplitude study and vortex structures

Firstly, note that in principle hydrodynamic friction can be reduced by simply increasing the clearance h_0, that is for Couette flow the decrease in wall shear stress will follow a $1/h_0$ law. However, in practice the clearance in the cylinder-piston system is required to be small for technical reasons so that a compromise is the use of periodically corrugated surface shapes with a fixed h_0 but varying amplitude a. Although the amplitude can be set to an arbitrarily high value this is in practice limited by the material properties of the components involved and manufacturing restrictions. Here the dependence of the friction coefficient κ on the amplitude is investigated which is expected to be more complex due to vortex generation in the associated valleys; the latter defines a relevant study object of its own due to its influence on fluid transport properties, temperature convection and related topics.

To begin with, Fig. 6.5 demonstrates the effect of increased amplitude of the surface contour on the streamline pattern in the case of Stokes flow; obviously at certain critical amplitudes a new eddy is born. This phenomenon is described in more detail in Fig. 6.6 (right), for increased amplitude the generation of new eddies (here, up to six) indicated by the black bullet-points; showing that for higher amplitudes, the critical amplitude and number of eddies are related essentially linearly and independent of the height h_0. To generate this plot roughly 70 simulations were run on the interval $a \in [0, 3\pi]$. Fig. 6.6 (left) is a plot of the scaled shear stress along the upper flat surface for selected amplitudes of the contoured surface, including those of Fig. 6.5. It shows that for small amplitude there is no eddy present; the shear stress exhibits only one maximum at $x = 0$. On formation of the first eddy two local maxima appear, the x-coordinates of which coincide with those of the streamline triple points; the value at $x = 0$ becomes a local minimum. In general, the number of eddies can be obtained from the number of local maxima of the surface shear stress curve divided by two as in Scholle [219].

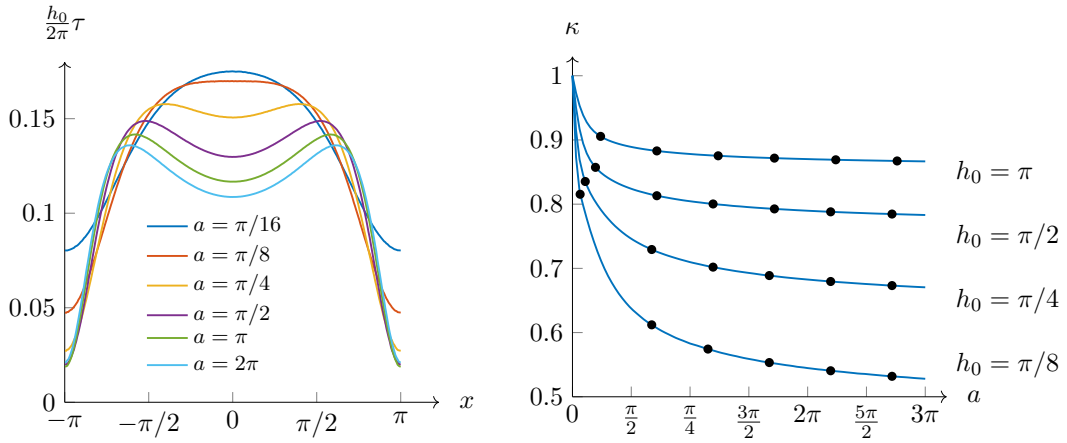


Figure 6.6: Scaled shear stress along the upper flat surface for the case of $h_0 = \pi/4$ and selected amplitudes a of the cosine-like surface contour (left). Critical amplitudes for the eddy formation in the presence of a cosine-like contour and under Stokes flow conditions (right); for selected height values h_0 the friction coefficient κ is plotted against the amplitude a while, for increasing amplitudes, the black bullet-points indicate the formation of a new eddy.

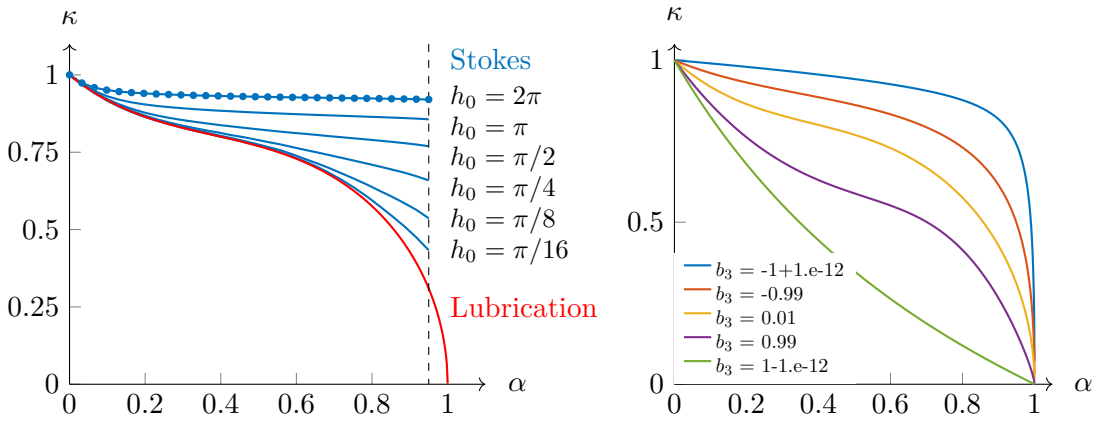


Figure 6.7: Friction coefficient κ dependence on the quotient $\alpha = \frac{a}{h}$ for a cosine-like contour (left): for selected height values h_0 the blue curves represent Stokes flow conditions while, independent of h_0, the red curve stems from the lubrication approximation; the results relate to those at 30 equidistant points (blue dots) on the interval $\alpha \in [0, 0.95]$. Contour study using the lubrication approximation (right): for selected parameters b_3 of the logarithmic surface contour $b(x)$, defined in paragraph (c) below, the friction coefficient κ is plotted against the quotient α; the shape of the curves does not depend on h_0.

In order to illustrate the friction behaviour of the co-sinusoidal corrugated system in the h_0-a-parameter domain in a simple way and to provide a clear comparison between lubrication theory and the Stokes solution it is instructive to introduce the quotient $\alpha :=$

$\frac{a}{h} = \frac{a}{a+h_0}$. Fig. 6.7 (left) plots the friction coefficient κ against α and for selected height values h_0 the blue curves result under Stokes flow conditions while, independent of h_0, the red curve results from the lubrication approximation. Obviously, $\alpha = 0$ corresponds to $a = 0$ and thus to the case of Couette flow with $\kappa = 1$; for $\alpha \to 1$, i.e. h_0 becomes small in comparison to a, generally friction decreases and in the limit case lubrication theory predicts zero friction. In the Stokes simulations the predicted κ-values increase with h_0 and altogether the possible Stokes curves are confined to the area between $\kappa = 1$ and the lubrication curve.

(c) Variation of the contour shape

The preceding investigations concentrated on the case of a cosine-like bottom contour which can indeed serve as a reasonable starting point. In the following it is investigated how a variation of the contour shape, depending on a preferably small number of geometric parameters, influences the hydrodynamic friction. For this purpose the following parametrised set of *logarithmic* surface contours is examined:

$$\begin{aligned} b(x) &= b_1 + b_2 \log(1 + b_3 \cos(x)) \,, \\ b_2 &= \frac{2a}{\log\left(\frac{1-b_3}{1+b_3}\right)} \,, \\ b_1 &= a - b_2 \log(1 - b_3) \,, \end{aligned}$$

depending on the geometry parameter $b_3 \in [-1, 1]$ only; the parameters b_1 and b_2 are chosen such that for a given amplitude a and for varying $b_3 \in [-1, 1]$ the identities $b(-\pi) = b(\pi) = a$ and $b(0) = -a$ always hold. The above contour in an elegant way of encompassing the transition from a concave pin contour ($b_3 \to -1$) to a convex box shape ($b_3 \to 1$) includes both extremes as special cases; the familiar cosine-like contour $b(x) = -a\cos(x)$ is obtained in the limit $b_3 \to 0$ as can be verified by L'Hôpital's rule. Fig. 6.8 shows streamline patterns for selected values of b_3 and under Stokes flow conditions, thus illustrating the transition process: taking $b_3 = 0$ as a starting point, as $b_3 \to -1$ an increasing number of eddies is generated, in the limit infinitely many; moving in the other direction, $b_3 \to 1$, there is just one eddy present until in the limit $b_3 = 1$ additional secondary corner eddies appear, cf. Moffatt [180].

For values of b_3 commensurate with those depicted in Fig. 6.8, Fig. 6.7 (right) shows the corresponding κ-α dependencies based on the lubrication model, giving when $b_3 = 0$ the case for a cosine-like contour, cf. Fig. 6.7 (left). As in the latter the corresponding Stokes curves for varying h_0 (but not depicted here) will reside in the area confined by the lubrication curves and the constant line $\kappa = 1$. Obviously the friction behaviour improves

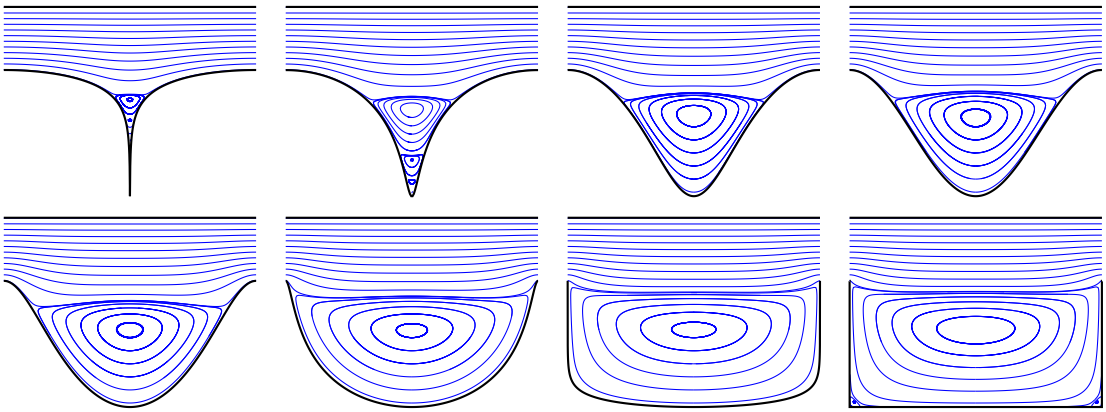


Figure 6.8: Vortex structure depending on the parameter b_3 of the logarithmic surface contour under Stokes flow conditions. The streamlines depicted above, chronologically from top left to bottom right, correspond to: $b_3 = -1 + 10^{-6}$, -0.99, -0.6, 0.01, 0.6, 0.99, $1 - 10^{-12}$ and 1. In the limit case $\lim_{b_3 \to -1}$ the contour converges to a needle point at $x = 0$, which is of no physical interest, unlike the limit case $\lim_{b_3 \to 1}$ which results in a box shape (below right) and is of interest.

for increasing b_3, which is less a contour inherent property but rather expressing the fact that friction decreases with the average gap width between the two surfaces. In the latter sense the box shaped contour is the logical optimum while a practically feasible contour, considering manufacturing restrictions and stability requirements for the ridges, will be much closer to $b_3 = 0$.

Fig. 6.9 illustrates, for fixed a and selected height values h_0, the explicit dependence of the friction coefficient κ on the geometry parameter b_3 of the logarithmic contour $b(x)$. For a simpler representation b_3 has been re-expressed via the parameter $c \in [-16, 16]$ and:

$$b_3 = \operatorname{sgn}(c) \cdot \left[1 - 10^{-|c|}\right].$$

The deviation between lubrication and Stokes solutions increases for $c > 0$ and thus the lubrication approximation overestimates the possible friction reduction significantly. The left graph suggests the parameter $c \in [0, 5]$, to be more or less independent of a and h_0; beyond the value $c = 5$, both solutions predict no significant further reduction in friction. The corresponding contours for $c = 1, \ldots, 4$ are depicted on the right-hand side emphasising the relevant effect: the contour peaks become slimmer and slimmer.

As mentioned before the hydrodynamic friction in the lubrication gap mainly depends on the average gap width, that is on the fluid volume between the upper and lower surface. In order to separate this effect out a study with constant volume is employed which is achieved by modifying the amplitude $a = a(b_3)$ depending on the parameter b_3 according

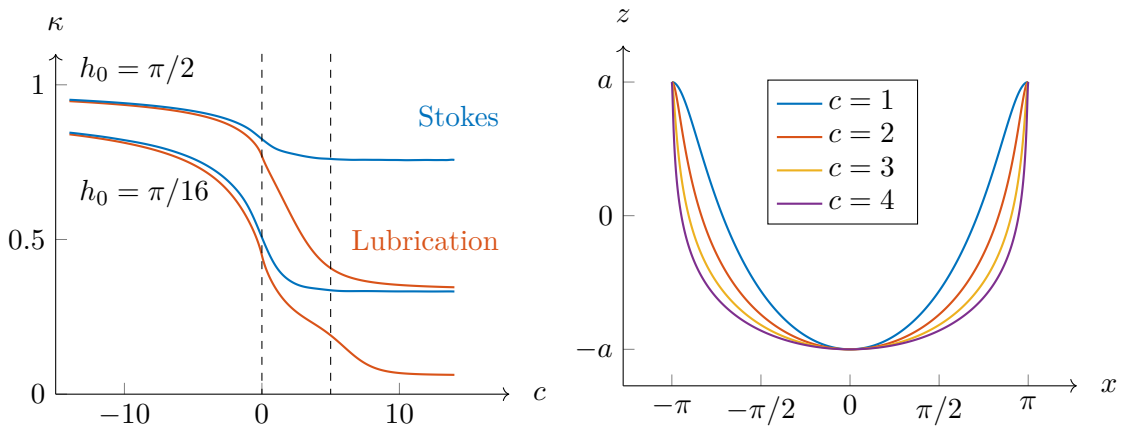


Figure 6.9: Friction coefficient κ dependence on the parameter c of the logarithmic contour (left). For $a = \pi/2$ and selected height values $h_0 = \pi/2$, $\pi/16$ the lubrication solution is given by the red and the FE Stokes solution by the blue curves. The corresponding contours for $c = 1, \ldots, 4$ are depicted on the right-hand side.

to the formula:

$$v(a; b_3) = \int_{-\pi}^{\pi} a - b(x)\,\mathrm{d}x \stackrel{!}{=} V = \text{constant}\,,$$

thus giving:

$$\int_{-\pi}^{\pi} b_2 \log(1 - b_3) - b_2 \log(1 + b_3 \cos(x))\,\mathrm{d}x = V\,,$$

and:

$$b_2(b_3) = \frac{1}{V}\left[2\pi \log(1 - b_3) - \int_{-\pi}^{\pi} \log(1 + b_3 \cos(x))\,\mathrm{d}x\right]\,, \tag{6.8}$$

$$a(b_3) = \frac{b_2}{2} \log\left(\frac{1 - b_3}{1 + b_3}\right)\,, \tag{6.9}$$

$$b_1(b_3) = a - b_2 \log(1 - b_3)\,. \tag{6.10}$$

The left-hand side of Fig. 6.10 displays the friction coefficient κ as a function of the dimensionless volume $v(a = \pi/2, b_3)$ of the logarithmic surface contour; $V = 4\pi a = 2\pi^2$ denotes the maximum possible volume for a given amplitude a. For $a = \pi/2$ and selected height values $h_0 = \pi/2$ and $h_0 = \pi/16$, the lubrication approximation results in the red and the FE Stokes solution in the blue curves. The dependency of the Stokes friction coefficient on the volume v is nearly linear while lubrication theory predicts a significant deviation from a linear profile when approaching a box shaped contour. In order to obtain a pure contour driven effect, on the right-hand side of Fig. 6.10 the volume impact has

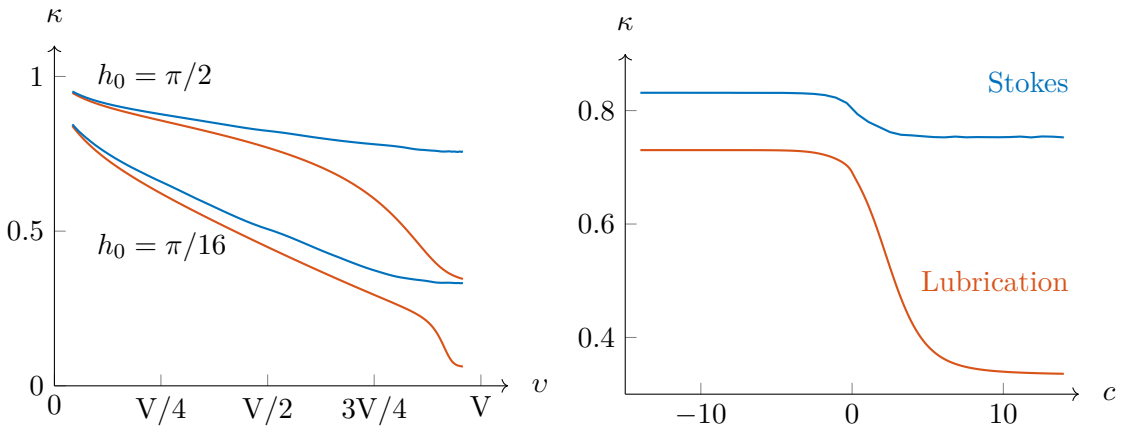


Figure 6.10: Friction coefficient κ dependence on the dimensionless volume $\upsilon(a = \pi/2, b_3)$ of the logarithmic contour, the latter being denoted in terms of fractions of the maximal possible volume $V = 4\pi a$ (left); for $a = \pi/2$ and selected height values $h_0 = \pi/2, \pi/16$ the lubrication approximation leads to the red and the FE Stokes solution to the blue curves. For $h_0 = \pi/2$ and constant $\upsilon(a(b_3), b_3) = V = 2\pi^2$ the κ-values of the lubrication (red) and FE Stokes (blue) solution are displayed versus the contour parameter c (right); the amplitudes are adapted according to (6.9).

been separated out: the amplitude of the logarithmic contour has been adapted according to (6.9) such that the volume $\upsilon(a(b_3), b_3) = V$ always remains constant. For $h_0 = \pi/2$ and a constant volume $V = 2\pi^2$ again the lubrication approximation and the FE Stokes solution result in similar red and blue curves as in Fig. 6.9.

(d) Inertial effects

Although inertial effects play a quantitatively minor role for the present problem, even for Reynolds numbers up to Re = 100, it is useful to obtain an overview of the basic mechanisms and how the eddy structure and the friction coefficient are influenced by an increase of the Reynolds number.

Fig. 6.11 illustrates the evolution of streamline patterns for increasing Reynolds number up to Re = 60 in which the top row refers to the cosine-like contour ($c = 0$) and the second row to the logarithmic contour using $c = 2$. The geometric aspect ratio in all cases is given by $h_0 = \frac{3}{2}a = \frac{\pi}{4}$, while for visualisation a different but for all cases identical scaling has been used. For the logarithmic contour inertial effects on the streamlines are, along with symmetry breaking, much more pronounced; this is also reflected in the variation of the friction coefficient: in the case of the cosine-like contour and Re = 60 a 1.9% increase in κ results compared to the Stokes case while for the logarithmic contour an approximate 7.7% increase is observed. Generally speaking, the advantage of a higher average gap for

large b_3 compared to those contours with $b_3 \approx 0$ seems to diminish for increased Reynolds numbers; the sharper ridge contours cause stronger vortex activity and asymmetry which again leads to increased friction. It would be worth extending investigation for higher Reynolds numbers, obviously at least $\mathrm{Re} > 100$ which is out of scope of the present study, as to whether a trade off point can be found.

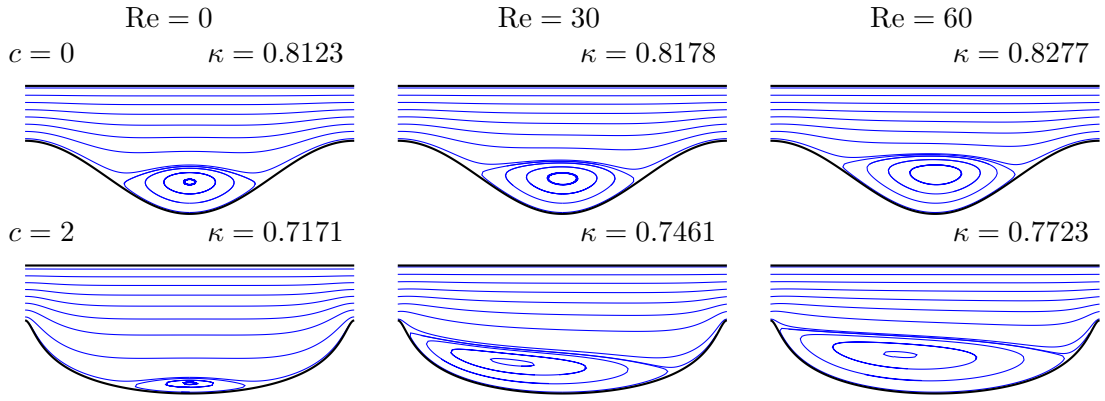


Figure 6.11: Evolution of streamline patterns for $\mathrm{Re} = 0$, 30, 60 (from left to right) for the case of the cosine-like contour using $c = 0$ (top row) and the logarithmic contour using $c = 2$ (bottom row). The geometric aspect ratio in all cases is $h_0 = \frac{3}{2}a = \frac{\pi}{4}$, while for the visualisation a different but for all cases identical scaling has been used.

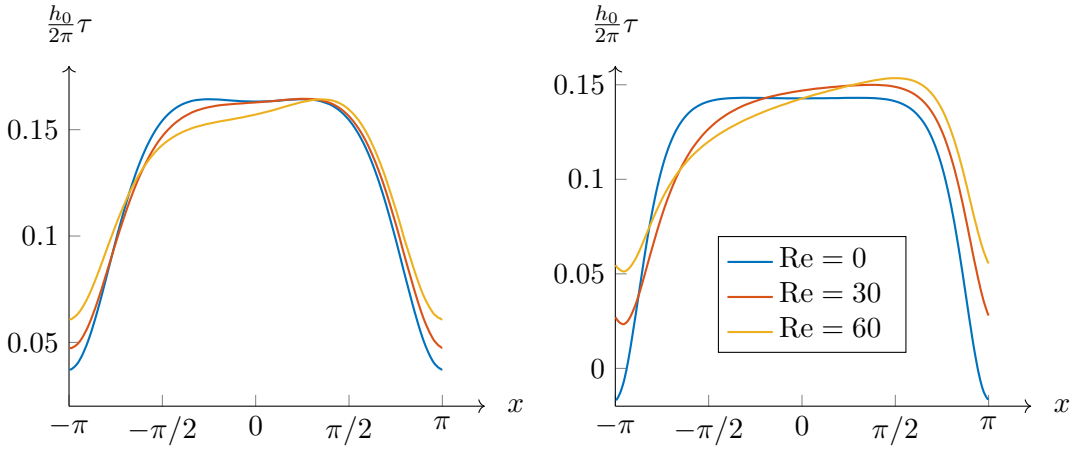


Figure 6.12: Scaled shear stress along the upper surface for $h_0 = \frac{3}{2}a = \pi/4$ and $\mathrm{Re} = 0$, 30, 60. The shear stresses correspond to the streamline patterns in Fig. 6.11, for $c = 0$ (left) and $c = 2$ (right).

Fig. 6.12 displays corresponding scaled shear stress distributions along the upper surface for $h_0 = \frac{3}{2}a = \pi/4$ and $\mathrm{Re} = 0,\ 30,\ 60$: on the left-hand side for the cosine-like contour and on the right-hand side for the logarithmic contour. With increasing Reynolds numbers the asymmetry of the velocity field becomes more pronounced which is also reflected in the shear stress curve. Qualitatively the shear stresses for both contours turn out to be similar. Fig. 6.13 displays the friction coefficient κ depending on the Reynolds number $\mathrm{Re} = \rho U_0 \lambda/(2\pi\eta)$ for the three different contour parameters $c = -2, 0, 4$ (from left to right) in the case of a logarithmic contour. For fixed amplitude $a = \pi/2$, three height values are considered: $h_0 = \pi$ (yellow), $h_0 = \pi/2$ (blue) and $h_0 = \pi/8$ (red). The friction coefficient increases with the Reynolds number, the effect for small Re being only moderate; with decreasing height h_0 the relevance of non-linear effects increases significantly.

Fig. 6.14 displays the friction coefficient κ depending on the quotient $\alpha = \frac{a}{h}$ for the cosine-like contour. For selected height values h_0 the blue curves result for Stokes flow ($\mathrm{Re} = 0$), the green curves for NS flow with $\mathrm{Re} = 75$ and the red curves depict the corresponding lubrication approximation (independent of h_0). This figure complements Fig. 6.7 (left) by including inertial effects thus combining the dependency of κ on α and h_0 for the three different flow regimes in just one picture.

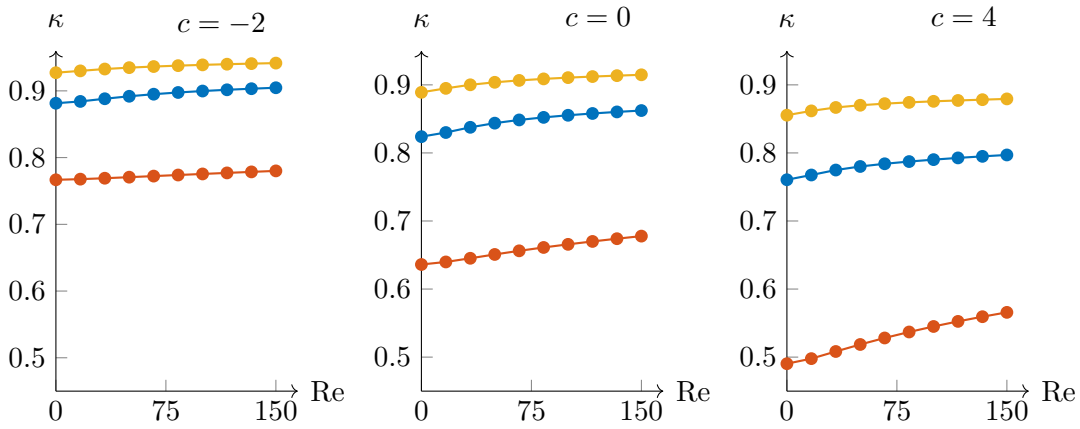


Figure 6.13: Friction coefficient κ depending on Re for the three different contour parameters $c = -2, 0, 4$ (from left to right); for a fixed amplitude $a = \lambda/4$, three height values are considered: $h_0 = \pi$ (yellow), $h_0 = \pi/2$ (blue) and $h_0 = \pi/8$ (red).

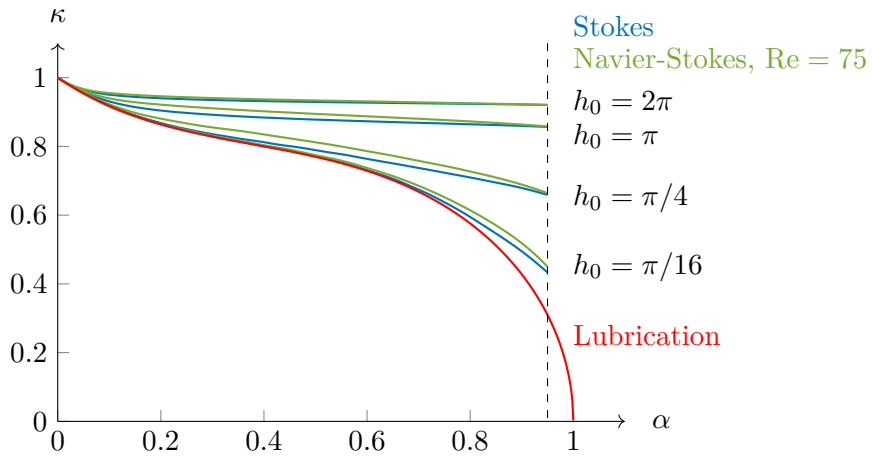


Figure 6.14: Friction coefficient κ depending on the quotient $\alpha = \frac{a}{h}$ for the cosine-like contour. For selected height values h_0 the blue curves result under Stokes conditions (Re = 0), the green curves under NS conditions with Re = 75 and the red curve depicts the lubrication approximation (independent of h_0). For this diagram the flow field has been evaluated for 30 equidistant points on the interval $\alpha \in [0, 0.95]$.

6.2 Gravity-driven steady film flow over periodic topography

Over the past two decades, considerable interest has been generated in the motion of liquid films over rigid surfaces containing well defined topographical features, mainly because of their importance in many branches of technology, such as functional thin film coatings [147, 274]. Relevant applications are manifold, comprising diverse fields like protective coatings in the ceramic and metalworking industry [139], the production of photo films and similar multilayer materials [147], high precision coatings in the microelectronics and semiconductor industry, falling film heat exchangers [268, 273], optical coatings and many more. To some extent prototypic film flow models cover macroscopic geological phenomena such as avalanches [165], glaciers [112] and debris flows [132] as well.

The primary motivation for considering topological features is that surfaces encountered in practice are never completely flat; rather, they contain complicating features such as well defined irregularities in the form of sharp steps or trenches or regular corrugations arising from nature or as a consequence of a particular manufacturing process. Despite the comparably simple geometric configuration gravity-driven film flow exhibits a wide range of complex physical phenomena like free-surface wave evolution [78], eddy genesis [228], film break-up [147] or rivulet [238, 239] and droplet formation [158]. Understanding the impact of various disturbances on flow structure is necessary to control and optimize relevant characteristics like the surface shape and stability [7, 39, 193, 234, 267], the

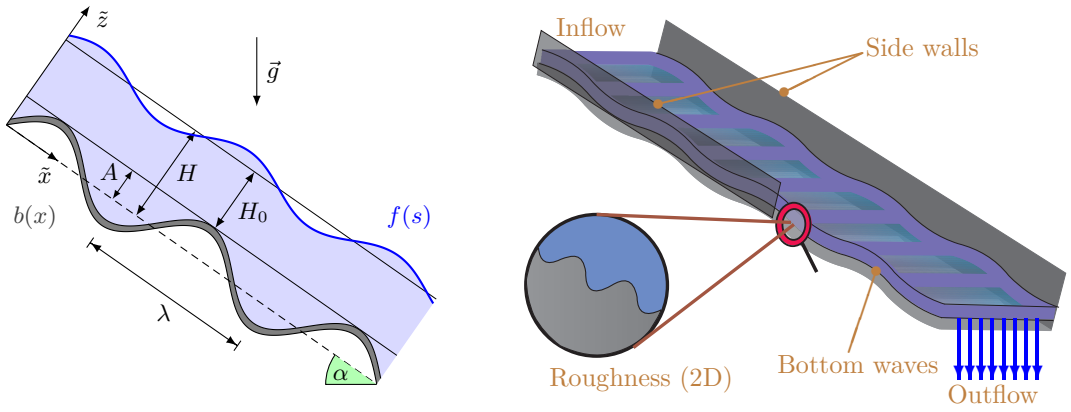


Figure 6.15: Gravity-driven film flow down an inclined corrugated rigid substrate in an experimental set-up (right) [219] and as an idealised two-dimensional model (left).

evolution of the vortex profile [228, 278] and the related mass transport [279] as well as characteristics influencing the temperature distribution and the drying process [147].

In the following the case of two-dimensional gravity-driven film flow over periodically corrugated topography inclined at an angle of α to the horizontal [228, 257, 258, 266, 267] is considered as shown in Fig. 6.15. Investigations focus on the computation of steady and periodic solutions of the full 2D Navier-Stokes equations involving inertial and capillary effects which generally exist under restricted conditions only [277], i.e. low Reynolds numbers depending on the inclination angle. Experimental results are known from [234, 278] indicating that steady solutions can, under certain conditions, also exist for higher Reynolds numbers. Normally in the direction of increasing Reynolds number a kind of pseudo-stable state is reached exhibiting a low-amplitude periodic movement of the free surface which leaves the eddy structure of the internal flow essentially the same [117]; thus, when the focus lies on the inner flow structure, it can be helpful to compute steady-state solutions which do not exist physically.

The periodic film flow problem is tackled by two different methods of solution: first, a spectral Fourier method for Stokes flow based on the complex-valued form of the first integral as described in Sec. 3.2; second, the LSFEM approach also including inertial effects based on the considerations in Chapter 5 and particularly Sec. 5.1.2.

6.2.1 Spectral complex-valued solution representation

An efficient procedure for solving free surface problems is given by the least-squares FE method of Chapter 5, particularly for the periodic flow described in Sec. 5.1.2. However, especially for the latter case one would usually prefer a Fourier basis instead of locally supported basis functions which would be suitable for stability analysis as well. This is the reason why, in the following, an alternative discretisation is proposed to solve for the

special case of steady, gravity-driven film flow over periodically corrugated substrates. Considerations are confined to the Stokes flow case although a generalisation to $\mathrm{Re} > 0$ is in principle possible. Scholle [219] provides an overview of how such a problem is tackled by means of the Stokes equations in primitive variables both for the limit case of thick films assuming a flat surface and for thin films allowing for a wavy surface.

Consider the 2D first integral of NS equations (3.12), (3.13) in terms of complex coordinates $\xi := x + \mathrm{i}z$ and its conjugate $\overline{\xi}$. In the steady Stokes flow case, (3.13) is a scalar equation for the determination of the pressure, which is of no importance here, while equation (3.12) reduces to a bianalytic equation $\partial^2\chi/\partial\overline{\xi}^2 = 0$ for the complex-valued potential field $\chi = \Phi + \mathrm{i}\eta\Psi$; the real part of which is the scalar potential field Φ and the imaginary part essentially the streamfunction Ψ. The scaling is similar to that of Sec. 6.1.2:

$$x = \frac{2\pi}{\lambda}\tilde{x}\,, \quad z = \frac{2\pi}{\lambda}\tilde{z}\,, \quad u = \frac{\tilde{u}}{u_n}\,, \quad \Psi = \frac{2\pi\tilde{\Psi}}{\lambda u_n}\,, \quad \Phi = \frac{2\pi\tilde{\Phi}}{\lambda\eta u_n}\,, \tag{6.11}$$

using as characteristic quantities the wavelength λ and the Nusselt velocity:

$$u_n = \frac{\varrho g\lambda^2 \sin\alpha}{8\pi^2\eta}\,; \tag{6.12}$$

and leading to a Reynolds number of $\mathrm{Re} = \varrho u_n\lambda/(2\pi\eta)$. The physical quantities are then denoted by capital letters and dimensionless quantities by small letters, analogous to Fig. 6.2 in Sec. 6.1.2. The domain is confined to $x \in [-\pi, \pi]$ and the solution is assumed 2π-periodic.

As mentioned in Secs. 1.2.1(b) and 3.1 the general solution of the bianalytic equation can be written in terms of two analytic functions $P(\xi)$ and $Q(\xi)$ which, without loss of generality, can be modified to $\chi = \Phi_n + \mathrm{i}(\Psi_n + P(\xi) + \overline{\xi}Q(\xi))$[1], thus separating the Nusselt solution $\Phi_n + \mathrm{i}\Psi_n$ from the rest. A representation for the streamfunction then results as:

$$\Psi = \mathrm{Im}(\chi) = hz^2 - \frac{z^3}{3} + P(\xi) + \overline{P(\xi)} + z\left(Q(\xi) + \overline{Q(\xi)}\right)\,, \tag{6.13}$$

with the non-dimensional height $h = 2\pi H/\lambda$; in equation (6.13) the x-dependant part, stemming from $\overline{\xi}$, has been omitted due to the assumed 2π-periodicity of the streamfunction and the result coinciding with that obtained by Scholle [219]. Moreover, the following derivative of Φ is used:

$$\begin{aligned} 2\frac{\partial\Phi}{\partial\overline{\xi}} = 2\frac{\partial}{\partial\overline{\xi}}\,\mathrm{Re}(\chi) = \mathrm{i}x(h-z) - \frac{x^2+z^2}{2} + hz - p_0(x+\mathrm{i}z) \\ + \overline{Q(\xi)} - Q(\xi) - 2\mathrm{i}\left(\overline{P'(\xi)} + z\overline{Q'(\xi)}\right)\,, \end{aligned} \tag{6.14}$$

[1] Multiplication with i and separation of the Nusselt solution is performed to keep in track with the considerations in [219] and is done without loss of generality.

in which the non-dimensioned Nusselt solution Φ_n plus derivatives have been taken from (5.42)-(5.44).

6.2.2 Boundary conditions in terms of analytic functions

In what follows the interest lies in obtaining the streamfunction according to (6.13); subsequently the velocities can be computed as a byproduct. In order to fix the degrees of freedom provided by the analytic functions $Q(\xi)$ and $P(\xi)$, the streamfunction representation has to fulfil the necessary boundary conditions, cf. Fig. 6.15 (left), in this case wall boundary conditions at the lower rigid surface and free boundary conditions at the upper free surface; the periodicity of the solution will be guaranteed by the choice of Q and P.

(a) Substrate boundary conditions

No-slip and no-penetration boundary conditions at the substrate contour $z = b(x)$ are required which, according to the streamfunction definition and (6.13), can be written as:

$$0 = u = -2\mathrm{i}\frac{\partial \Psi}{\partial \overline{\xi}} = 2hz - z^2 + Q(\xi) + \overline{Q(\xi)} - 2i\left(\overline{P'(\xi)} + z\overline{Q'(\xi)}\right), \tag{6.15}$$

for $z = b(x)$ and $\xi = x + \mathrm{i}b(x)$.

(b) Kinematic BC at the free surface

Let the free surface $(x(s), z(s)) = (f_1(s), f_2(s))$ be parametrised with respect to arclength s. Normal and tangential unit vectors are then given by $t_i(s) = f_i'(s)$, $n_i(s) = \varepsilon_{ji} t_j(s)$ as in Sec. 3.2.2. A complex-valued description is provided by $f(s) = f_1(s) + \mathrm{i}f_2(s)$, $t = f'(s)$, $n = \mathrm{i}f'(s)$. For convenience the shape of the free surface is represented by the derivative of a periodic function g:

$$z = h + g'(x)\,. \tag{6.16}$$

In the case of stationary flow the kinematic boundary condition is equivalent to identifying the free surface as a streamline, i.e. $\Psi(f) = \psi_s = \text{const.}$, which leads, after substitution of (6.16) into formula (6.13), to the final form:

$$\psi_s = \frac{2}{3}h^3 + h^2 g'(x) - \frac{g'(x)^3}{3} + 2\mathrm{Re}(P(\xi) + zQ(\xi))\,, \tag{6.17}$$

for all $x \in [-\pi, \pi]$ with $z = z(x) = h + g'(x)$ and $\xi = \xi(x) = x + \mathrm{i}z(x)$.

(c) Dynamic BC at the free surface

In the case of stationary Stokes flow with constant surface tension the dynamic boundary condition (3.23) in dimensionless form reduces to:

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\hat{\sigma}f' + 4\mathrm{i}\frac{\partial\Phi}{\partial\overline{\xi}}\right) = Un\,, \tag{6.18}$$

with the following dimensionless quantities: when $\tilde{U}(\tilde{x},\tilde{z}) = \varrho g[\cos(\alpha)\tilde{z} - \sin(\alpha)\tilde{x}]$ denotes the potential energy density in dimensional form, the scaling (6.11) leads to:

$$U(x,z) = \frac{\tilde{U}\lambda}{2\pi\eta u_n} = 2[\cot(\alpha)z - x]\,; \tag{6.19}$$

accordingly the Capillary number is given by $\mathrm{Ca} = \eta u_n/\sigma$, but for convenience the inverse $\hat{\sigma} := \mathrm{Ca}^{-1}$ is used.

In equation (6.18) the term involving the potential energy density can, by use of the identity $n = \mathrm{i}x' - z'$, be rewritten as:

$$Un = 2(z\cot\alpha - x)n = \frac{\mathrm{d}}{\mathrm{d}s}\left(2xz - \mathrm{i}x^2 - \cot(\alpha)z^2\right) + 2\left(\mathrm{i}\cot\alpha - 1\right)zx'\,, \tag{6.20}$$

$$= \frac{\mathrm{d}}{\mathrm{d}s}\left(2xz - \mathrm{i}x^2 - \cot(\alpha)z^2 + 2(\mathrm{i}\cot\alpha - 1)(hx + g(x))\right)\,, \tag{6.21}$$

in which zx' has, via (6.16), been replaced by the more convenient expression:

$$zx' = (h + g'(x))x'(s) = \frac{\mathrm{d}}{\mathrm{d}s}\left(hx + g(x)\right)\,.$$

Substitution of (6.14) and (6.21) into the dynamic boundary condition (6.18) gives:

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}s}\Big(&\hat{\sigma}f' + 2\left(1 - \mathrm{i}\cot\alpha\right)g + \left(\cot\alpha - \mathrm{i}\right)z^2 + 2i\left(\bar{Q} - Q\right) + 4\left(\overline{P'} + z\overline{Q'}\right) \\ &- 2h\cot(\alpha)z + 2ihz\Big)\,, \end{aligned} \tag{6.22}$$

in which the pressure constant in (6.14) has been set to $p_0 = -h\cot\alpha$ in order to eliminate remaining x-dependencies. Integration of (6.22) and substitution of (6.16) then results in:

$$\begin{aligned} C = {}&\hat{\sigma}f' + 2\left(1 - \mathrm{i}\cot\alpha\right)g - h^2\cot\alpha + \cot(\alpha)g'^2 + \mathrm{i}h^2 - \mathrm{i}g'^2 \\ &+ 2\mathrm{i}\left(\bar{Q} - Q\right) + 4\left(\overline{P'} + z\overline{Q'}\right)\,, \end{aligned}$$

involving an arbitrary integration constant C; the choice of $C := \hat{\sigma} - (\cot\alpha - \mathrm{i})h^2$ delivers:

$$0 = \hat{\sigma}(f' - 1) + 2\left(1 - \mathrm{i}\cot\alpha\right)g + \cot(\alpha)g'^2 - \mathrm{i}g'^2 + 2\mathrm{i}\left(\bar{Q} - Q\right) + 4\left(\overline{P'} + z\overline{Q'}\right)\,. \tag{6.23}$$

The following simple relations will be utilised subsequently:

$$2\mathrm{i}(\overline{Q}-Q)\overline{f'}+4(\overline{P'}+z\overline{Q'})\overline{f'}=4\frac{\mathrm{d}}{\mathrm{d}s}(\overline{P}+z\overline{Q})+4\,\mathrm{Im}\left(Q\overline{f'}\right)\,,$$
$$\overline{f'}=x'-\mathrm{i}z'=x'-\mathrm{i}(g''x')=\left(1-\mathrm{i}g''\right)x'\,,$$
$$1=f'\overline{f'}=\sqrt{x'(s)^2+z'(s)^2}=\sqrt{1+g''(x)^2}x'(s)\,.$$

Multiplication of equation (6.23) by $\overline{f'}$ and splitting of the resulting product into real and imaginary part yields, under consideration of the above relations, two equations:

$$0=\frac{\hat{\sigma}}{2}\left(\sqrt{1+g''^2}-1\right)+\left(1-\cot(\alpha)g''\right)g+\frac{\cot\alpha}{2}g'^2-\frac{g'^2g''}{2}$$
$$+2\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{Re}\left(P+zQ\right)+2\,\mathrm{Im}\left(Q(1-\mathrm{i}g'')\right)\,, \tag{6.24}$$
$$0=\frac{\hat{\sigma}}{2}g''-(g''+\cot\alpha)g-\frac{\cot\alpha}{2}g'^2g''-\frac{g'^2}{2}-2\frac{\mathrm{d}}{\mathrm{d}s}\,\mathrm{Im}\left(P+zQ\right)\,. \tag{6.25}$$

Inserting the derivative of (6.17),

$$2\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{Re}(P(\xi)+zQ(\xi))=\left(g'^2-h^2\right)g''\,,$$

into (6.24) leads to:

$$\hat{\sigma}\frac{\sqrt{1+g''^2}-1}{2}+g-h^2g''+\frac{g'^2g''}{2}+\cot\alpha\left(\frac{g'^2}{2}-gg''\right)+2\,\mathrm{Im}\left(Q(1-ig'')\right)=0\,. \tag{6.26}$$

The integral of (6.25) with respect to x:

$$\frac{\hat{\sigma}}{2}g'-\int\left(gg''+\frac{g'^2}{2}+\cot(\alpha)g\right)\mathrm{d}x-\frac{\cot\alpha}{6}g'^3-2\,\mathrm{Im}\left(P+zQ\right)=0\,, \tag{6.27}$$

can, in combination with the kinematic boundary condition (6.17), conveniently be compounded to form a complex-valued "hybrid" boundary condition:

$$\left(h^2-\frac{\mathrm{i}}{2}\hat{\sigma}\right)g'+\mathrm{i}\int\left(gg''+\frac{g'^2}{2}+\cot\alpha g\right)\mathrm{d}x+\frac{\mathrm{i}\cot\alpha-2}{6}g'^3+2(P+zQ)=\psi_s-\frac{2}{3}h^3\,. \tag{6.28}$$

6.2.3 Fourier series discretisation

In order to discretise the wall condition (6.15) and the two free surface conditions (6.26), (6.28) the unknown functions $Q(\xi)$ and $P(\xi)$ are expanded as Fourier series:

$$Q(\xi)=\sum_{\substack{k=-\infty\\k\neq0}}^{\infty}Q_ke^{\mathrm{i}k\xi}+\frac{u_0+\mathrm{i}w_0}{2}\,,\qquad P(\xi)=\sum_{\substack{k=-\infty\\k\neq0}}^{\infty}\frac{Q_k-R_k}{2k}e^{\mathrm{i}k\xi}+\frac{1}{2}\psi_0\,, \tag{6.29}$$

depending on two series of complex-valued coefficients Q_k and R_k.

(a) Substrate boundary condition

Substitution of (6.29) into the substrate boundary condition (6.15) delivers an equation $0 = u_b(x)$ of the form:

$$0 = u_0 + 2hb(x) - b^2(x) + \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} e^{-kb(x)} \left(Q_k e^{\mathrm{i}kx} + \left(\bar{R}_k - 2kb(x)\bar{Q}_k \right) e^{-\mathrm{i}kx} \right) . \tag{6.30}$$

In principle, a system of infinitely many linear equations for the coefficients Q_k and R_k can be obtained from (6.30) by using the standard inner product in the following way:

$$(u_b(x), e^{-\mathrm{i}mx})_0 := \frac{1}{2\pi} \int_{-\pi}^{\pi} u_b(x) e^{-\mathrm{i}mx} \, \mathrm{d}x = 0 \,, \quad \text{for } m \in \mathbb{Z} \,; \tag{6.31}$$

for the case of a sinusoidal substrate contour $b(x) = -a\cos x = -\frac{a}{2}\left(e^{\mathrm{i}x} + e^{-\mathrm{i}x}\right)$, which is of particular interest here, Scholle [219] demonstrates that the linear system (6.31) can be reduced further to the more convenient form:

$$\delta_{0m} u_0 + \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \left(I_{m-k}(ka) Q_k + ka\big(I_{m+k-1}(ka) + I_{m+k+1}(ka)\big)\bar{Q}_k + I_{m+k}\,(ka)\,\bar{R}_k \right) = B_m \,, \tag{6.32}$$

with inhomogeneity:

$$B_m = \frac{a^2}{2}\delta_{0m} + ah\left(\delta_{1m} + \delta_{-1m}\right) + \frac{a^2}{4}\left(\delta_{2m} + \delta_{-2m}\right) \,, \tag{6.33}$$

involving the Kronecker delta, δ, and the modified Bessel functions of the first kind [1]:

$$I_n(x) := \frac{1}{\pi} \int_0^{\pi} e^{x\cos(\theta)} \cos(n\theta) \, \mathrm{d}\theta \,. \tag{6.34}$$

(b) Free surface conditions

Now, consider the two conditions (6.26), (6.28) at the free surface. These require the Fourier expansion of the surface function $g(x)$ and its derivatives:

$$g(x) = \sum_{k=-\infty}^{\infty} g_k e^{\mathrm{i}kx} \,, \quad g'(x) = \mathrm{i} \sum_{k=-\infty}^{\infty} k g_k e^{\mathrm{i}kx} \,, \quad g''(x) = -\sum_{k=-\infty}^{\infty} k^2 g_k e^{\mathrm{i}kx} \,, \tag{6.35}$$

along with $g_{-k} = \bar{g}_k$ due to g being real-valued. The Fourier expansions (6.29) of $Q(\xi)$ and $P(\xi)$ with $\xi = x + \mathrm{i}(h + g'(x))$ involve the inconvenient term $e^{-kg'(x)}$; if this term is approximated by a truncated Taylor series:

$$e^{-kg'(x)} \approx 1 - kg(x)' = 1 - \mathrm{i}k \sum_{l=-\infty}^{\infty} l g_l e^{\mathrm{i}lx} \,, \tag{6.36}$$

representations for Q and P at the free surface result as:

$$Q(x+\mathrm{i}(h+g'(x))) = \sum_{\substack{k=-\infty\\k\neq 0}}^{\infty} e^{-kh}Q_k\left(1-\mathrm{i}k\sum_{l=-\infty}^{\infty} lg_l e^{\mathrm{i}lx}\right)e^{\mathrm{i}kx} + \frac{u_0+\mathrm{i}w_0}{2}$$

$$= \sum_{\substack{k=-\infty\\k\neq 0}}^{\infty} e^{k(\mathrm{i}x-h)}Q_k - \mathrm{i}\sum_{\substack{k,l=-\infty\\k\neq 0}}^{\infty} k(l-k)g_{l-k}e^{\mathrm{i}lx-kh}Q_k + \frac{u_0+\mathrm{i}w_0}{2}\,, \tag{6.37}$$

and:

$$P(x+\mathrm{i}(h+g'(x))) = \sum_{\substack{k=-\infty\\k\neq 0}}^{\infty} e^{-kh}\frac{Q_k-R_k}{2k}\left(1-\mathrm{i}k\sum_{l=-\infty}^{\infty} lg_l e^{\mathrm{i}lx}\right)e^{\mathrm{i}kx} + \frac{1}{2}\psi_0 \tag{6.38}$$

$$= \sum_{\substack{k=-\infty\\k\neq 0}}^{\infty} e^{k(\mathrm{i}x-h)}\frac{Q_k-R_k}{2k} - \mathrm{i}\sum_{\substack{k,l=-\infty\\k\neq 0}}^{\infty}(l-k)g_{l-k}e^{\mathrm{i}lx-kh}\frac{Q_k-R_k}{2} + \frac{1}{2}\psi_0\,, \tag{6.39}$$

in which an index shift $l \leftarrow l+k$ has been used to arrive at the second equality in both cases. In a next step the series representations (6.35)-(6.39) are applied to the dynamic boundary condition (6.26) and to the hybrid boundary condition (6.28) which is described in detail in [176] and summarised in Appendices D.1.1 and D.1.2. There it is also shown that an inner product approach, similar to that in (6.31) for the substrate boundary condition, leads after mathematical manipulations to the two systems of non-linear equations:

$$\begin{aligned} &w_0\delta_{0m} + \sum_{l=-\infty}^{\infty}\left(\frac{\hat{\sigma}}{4}l^2(m-l)^2 + \frac{\cot\alpha}{2}l(3l-m)\right)g_l g_{m-l}\\ &\quad + \left(1+h^2m^2+u_0m^2\right)g_m + \sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}(m-k)(m-2k)\left(e^{-kh}Q_k + e^{kh}\bar{Q}_{-k}\right)g_{m-k}\\ &\quad - \mathrm{i}\left(e^{-mh}Q_m - e^{mh}\bar{Q}_{-m}\right)(1-\delta_{0m}) + \frac{1}{2}\sum_{k,l=-\infty}^{\infty}k^2(l-k)(m-l)g_k g_{l-k}g_{m-l}\\ &\quad - \mathrm{i}\sum_{k,l=-\infty}^{\infty}k(l-k)(m-l)^2\left(e^{-kh}Q_k - e^{kh}\bar{Q}_{-k}\right)g_{l-k}g_{m-l} = 0\,, \end{aligned} \tag{6.40}$$

for $(e^{\mathrm{i}mx},(6.26))_0 = 0$ and:

$$\begin{aligned} &\left(\left(\mathrm{i}h^2 + \frac{\hat{\sigma}}{2} + \mathrm{i}u_0 - w_0\right)m + \frac{\cot\alpha}{m}(1-\delta_{0m})\right)g_m + (\psi_0 + hu_0 + \mathrm{i}hw_0 + \mathrm{i}C)\,\delta_{0m}\\ &\quad + e^{-mh}\frac{(1+2hm)Q_m - R_m}{m}(1-\delta_{0m}) + \mathrm{i}\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}e^{-kh}(m-k)g_{m-k}\big((k-2hk)Q_k + R_k\big) \end{aligned}$$

$$-\frac{1}{2}\sum_{l=-\infty}^{\infty}\frac{l(l+m)}{m}g_l g_{m-l}\,(1-\delta_{0m})+2\sum_{k,l=-\infty}^{\infty}e^{-kh}k(l-k)(m-l)g_{l-k}g_{m-l}Q_k$$
$$+\frac{\cot\alpha+2\mathrm{i}}{6}\sum_{k,l=-\infty}^{\infty}k(l-k)(m-l)g_k g_{l-k}g_{m-l}=\left(\psi_s-\frac{2}{3}h^3\right)\delta_{0m}\,, \tag{6.41}$$

for $(e^{\mathrm{i}mx},(6.28))_0 = 0$, both for arbitrary $m \in \mathbb{Z}$. Note that always m is used as an equation index, while k and l are used as variable indices.

(c) Determination of the surface function's zero-mode

In order to solve for the discretised boundary conditions the zero-mode g_0 of the surface function g has yet to be determined. For this purpose equation (6.28) is evaluated at $x = \pm\pi$. The difference of both results leads, due to the 2π-periodicity of g', P and Q, to:

$$\int_{-\pi}^{\pi}\left(gg''+\frac{g'^2}{2}+\cot(\alpha)g\right)\mathrm{d}x=\int_{-\pi}^{\pi}\left(-\frac{1}{2}g'^2+\cot(\alpha)g\right)\mathrm{d}x=0\,,$$

in which the second equality is due to partial integration of the gg'' term. Using (6.35) it follows that:

$$\int_{-\pi}^{\pi}\left(\frac{1}{2}\sum_{k,l=-\infty}^{\infty}l(k-l)g_l g_{k-l}e^{ikx}+\cot\alpha\sum_{k=-\infty}^{\infty}g_k e^{ikx}\right)\mathrm{d}x=0\,. \tag{6.42}$$

Furthermore for each $k \in \mathbb{Z}$ the property $(1, e^{\mathrm{i}kx})_0 = \delta_{0k}$ holds which when applied to equation (6.42) yields:

$$\frac{1}{2}\sum_{l=-\infty}^{\infty}-l^2 g_l g_{-l}+\cot\alpha g_0=0\,;$$

due to being real-valued the zero-mode of the surface function g after manipulation finally reads:

$$g_0=\tan\alpha\sum_{l=1}^{\infty}l^2 g_l\bar{g}_l\,. \tag{6.43}$$

6.2.4 Spectral iterative method of solution

In order to solve the system of coupled equations (6.32), (6.40) and (6.41) the infinite series involved have to be truncated. Thus the complex-valued coefficients Q_k and R_k are confined to an index range of $-N_n \le k \le N_p$ with $k \in \mathbb{Z}$ and $N_p, N_n \in \mathbb{N}_{>0}$, while coefficients outside of this range are set to zero. Correspondingly, replacing $g_{-k} = \bar{g}_k$ and g_0 by (6.43) the index range of g_k can be confined to $1 \le k \le N_g \in \mathbb{N}_{>0}$; in summary, considering u_0 and w_0, a number of $2(N_n + N_p) + N_g$ complex-valued and 2 real-valued unknown coefficients arise. In order to obtain a consistent system with an equal number

of equations and unknowns, equations (6.32) and (6.41) are considered for $m \in \mathbb{Z}$, $-N_n \leq m \leq N_p$, while (6.40) is considered for $1 \leq m \leq N_g$.

The resulting system of equations is solved in a decoupled way by an iterative solution strategy, cf. Fig. 6.16. Starting with zero values for the negative modes, i.e. for Q_k, R_k with $k < 0$, the following two steps are repeated until the norm of the difference between two consecutive solutions falls below a certain threshold value ε:

(1) Using given values for the negative modes Q_k, R_k, $k < 0$ from the previous iteration step, the discretised substrate boundary condition (6.32) with $-N_n \leq m \leq N_p$ is solved for u_0 and for the positive modes Q_k and R_k with $k > 0$. The corresponding linear system of equations is:

$$\delta_{0m} u_0 + \sum_{k=1}^{N_p} \Big(I_{m-k}(ka) Q_k + ka \big(I_{m+k-1}(ka) + I_{m+k+1}(ka) \big) \bar{Q}_k + I_{m+k}(ka)\, \bar{R}_k \Big) = B_m^*, \tag{6.44}$$

involving the modified inhomogeneity:

$$\begin{aligned} B_m^* = {} & \frac{a^2}{2} \delta_{0m} + ah \left(\delta_{1m} + \delta_{-1m} \right) + \frac{a^2}{4} \left(\delta_{2m} + \delta_{-2m} \right) \\ & - \sum_{k=-N_n}^{-1} \Big(I_{m-k}(ka) Q_k + ka \big(I_{m+k-1}(ka) + I_{m+k+1}(ka) \big) \bar{Q}_k + I_{m+k}(ka)\, \bar{R}_k \Big). \end{aligned} \tag{6.45}$$

(2) Using given values of u_0 and the positive modes Q_k, R_k, $k > 0$ from the previous iteration step, the discretised non-linear boundary conditions at the free surface, i.e. (6.40) with $1 \leq m \leq N_g$ and (6.41) with $-N_n \leq m \leq N_p$, are solved for w_0, g_k and the negative modes Q_k, R_k with $k < 0$. Linearisation is performed via Newton's method the details of which are omitted here.

Although the convergence is not analysed in detail, practice shows that only very few modes (below 10) and iteration steps (below 15) are necessary to obtain highly accurate results. However, this is only true for comparatively small substrate waviness. For increasing substrate amplitude a it is observed that the solution near the free surface is still accurate while the series solution does not converge in the vicinity of the fixed substrate. This leads to a possibly low resolution of eddy structures, a phenomenon which was similarly observed by Scholle [219], while the free surface is well captured for, otherwise problematic, extremely thin films. The latter feature makes it particularly suitable if only the free surface shape is of interest or, for instance, if a subsequent stability analysis is desired.

It is further mentioned that the equation splitting and correspondingly the iteration *direction* in the above method is designed in a particular way: if the fixed-wall equations

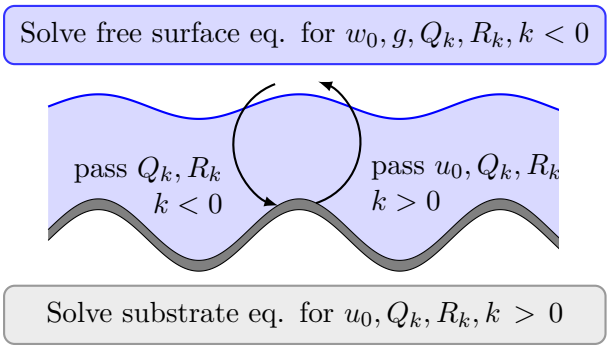


Figure 6.16: Iterative solution procedure as described by steps (1) and (2): the fixed-substrate equation (6.44) is solved for u_0 and the negative modes of Q, R which are then passed to the free-surface equations (6.40), (6.41). The latter are solved for w_0, g_k and the positive modes of Q, R which are again passed to the substrate equation etc.

are solved for the positive modes $k > 0$, the corresponding complex-valued exponential functions in (6.29) *decay* in the direction of the free surface, thus influencing the free surface equations only in a minor way. On the other hand the free surface equations are solved for negative modes $k < 0$ the exponential functions of which decay in wall direction. This approach ensures that the iteration remains stable which is probably not the case if the iteration is performed in the opposite direction. Moreover, from a numerical perspective, this can be interpreted as an unconventional way to circumvent the conditioning problem which is otherwise involved, an interesting perspective for future investigations.

6.2.5 Alternative solution by the LSFEM

The above spectral method is based on an analytic function representation which is confined to Stokes flow and, in the above implementation, also to sinusoidal contour shapes of the substrate. In the case of $\mathrm{Re} > 0$ and arbitrary substrate contours the first integral formulation of the full steady 2D-NS equations (4.12) is solved by the LSFEM of Chapter 5, the same as employed for the friction problem in Sec. 6.1, using in particular the procedure for film flows given in Sec. 5.1.2.

The boundary conditions are in principle the same as given for the spectral method, cf. also Fig. 5.1 of Sec. 5.1.2, but of course are treated differently in a FE context. The problem domain is assumed periodic on the left and right-hand side, which is assured by an identification of the respective grid points, while a pair of no-slip and no-penetration velocity conditions, given along the substrate, are treated as standard Dirichlet conditions. Moreover, three boundary conditions are required along the free surface, a kinematic condition (no flow penetration of the surface) and two dynamic conditions (stress equilibrium).

As two conditions are sufficient to determine a problem with fixed domain, the shape of the free surface can be found by iterating over one of the conditions while solving a sequence of flow problems with the other two conditions set on a fixed domain.

In this vein, the complete non-linear problem with variable surface shape is solved by the iterative procedure of Algorithm 2 presented in Sec. 5.1.2. In each iteration step a simpler problem with a fixed approximation of the free surface is solved only prescribing the dynamic BC, and between the iteration steps the surface is updated via the kinematic condition by solving equation (5.49). Each of the non-linear sub-problems is solved by Newton's method as demonstrated in Algorithm 1 relying on the matrix construction of Sec. 5.1.1 and the isoparametric concept of Sec. 5.1.3.

For reasons of clarity non-dimensionalisation of the first integral is specified for which the inclination angle, film height, substrate amplitude and wave length are defined as in Fig. 6.15. The relevant physical quantities are either denoted by a tilde or by capital letters similar to Sec. 6.1; in contrast to the scaling employed in (6.11), (6.12) and (6.19) which is particularly useful for the Fourier discretisation, it is more appropriate for thin films, and also for comparison with results in the literature, to base the scaling on the film height rather than on the wave length. Using the film height $H_0 = H - A$ as a characteristic length scale, the corresponding non-dimensional quantities are obtained as:

$$x = \frac{\tilde{x}}{H_0}\,, \qquad z = \frac{\tilde{z}}{H_0}\,, \qquad a = \frac{A}{H_0}\,, \qquad \ell = \frac{\lambda}{H_0}\,, \qquad h_0 = 1\,. \tag{6.46}$$

As a characteristic velocity the free surface Nusselt velocity $u_{n,h}$ of a film with height H_0 is assumed, i.e.:

$$u_{n,h} = \frac{\varrho g H_0^2 \sin(\alpha)}{2\eta}\,, \tag{6.47}$$

leading to the scaling of the velocities u_i, the potential fields ϕ_i and the potential energy density $\tilde{U}(\tilde{x},\tilde{y}) = \varrho g[\cos(\alpha)\tilde{y} - \sin(\alpha)\tilde{x}]$ as:

$$u_i = \frac{\tilde{u}_i}{u_n}\,, \qquad \phi_i = \frac{\tilde{\phi}_i}{\eta u_n}\,, \qquad U_h(x,y) = \frac{\tilde{U} H_0}{\eta u_n} = \frac{1}{2}[\cot(\alpha)y - x]\,,$$

for $i = 1, 2$. Then, with a Reynolds number $\mathrm{Re}_h = \varrho u_n H_0/\eta$ and a Capillary number $\mathrm{Ca}_h = \eta u_n/\sigma^2$ the non-dimensioned boundary value problems solved in each free surface iteration step, using the operator notation from Sec. 4.2 with $u = (u_1, u_2, \phi_1, \phi_2)$, are of the form:

[2] The index h of the non-dimensional numbers distinguishes them from the above used numbers based on a wave length scaling. However, it is easily seen that: $\mathrm{Ca}_h = \left(\frac{2\pi H_0}{\lambda}\right)^2 \mathrm{Ca}$ and $\mathrm{Re}_h = \left(\frac{2\pi H_0}{\lambda}\right)^3 \mathrm{Re}$.

$$\mathcal{L}u + \mathcal{N}[\mathrm{Re}_h, u]u = 0\,, \qquad \text{in the domain}\,, \tag{6.48a}$$

$$u_i(s) = 0\,, \qquad \text{along the bottom}\,, \tag{6.48b}$$

$$2\varepsilon_{ij}\phi_j(s) = \mathrm{Ca}_h^{-1} t_i(s) - \int_{s_0}^{s} U_h(\tilde{s}) n_i(\tilde{s})\,\mathrm{d}\tilde{s}\,, \qquad \text{along the free surface}\,. \tag{6.48c}$$

It is noted that the above kind of iterative splitting avoids the necessity to incorporate the geometric degrees of freedom into the system matrix, as in the case of the spine method [265], which would result in a more complicated implementation process and destroy the favourable structure of the occurring matrices[3]. In contrast the linear systems in each iteration step can be constructed and solved efficiently via the multigrid procedure of Sec. 5.4.3. On the downside the iterative procedure is heuristic and may in selected cases take over a hundred iteration steps; however, the number of surface iterations in the presence of a cosine-like substrate contour is usually found to be between 10 and 20. The step width is controlled by a complex adaptive relaxation scheme which can smooth the correction of the free surface automatically to avoid numerical instabilities; in severe cases also a Reynolds number stepping method is activated. As a convergence criterion the fraction between the flow in the normal direction of the surface and flow in the tangential direction is used which is required to fall below a value of 0.1%. The domain is approximated by a structured triangulation (6-point triangles) using continuous, piecewise quadratic basis functions for all test and solution spaces and the boundary triangles are allowed to be curved in accordance with the isoparametric concept. In the following study usually about 50,000 grid points are used. The mesh is only created once, while in the course of the iteration only the y-values of the coordinates are changed with the surface.

6.2.6 Numerical results

The aim of this section is less a parameter study on periodic film flow which has been extensively carried out by several authors [40, 186, 228, 266, 267], but to demonstrate that the first integral approach, both the above described spectral variant and the LSFEM, are suitable for and efficient at solving such problems. First, results obtained by the LSFEM are compared to the experimental results by Schörner et al. [234] and Wierschem et al. [278] in Figs. 6.17 and 6.18; they are found to be in very good agreement and serve as a verification of the FE methodology[4].

The parameter studies in Figs. 6.19 - 6.22 demonstrate that the first integral FE solver is capable of accurately solving a number of periodic film flow problems for variously shaped

[3] The splitting is also fundamentally different from the one for the spectral method described in Sec. 6.2.4.

[4] The calculated Reynolds and Capillary numbers deviate slightly from those in [278] who, in contrast to (6.46), (6.47), use $h_n = \sqrt[3]{3\nu\dot{q}/(g\sin(\alpha))}$ with flow rate $\dot{q}$ as characteristic length scale and the mean Nusselt velocity $u_n = \dot{q}/h_n$ as the characteristic velocity.

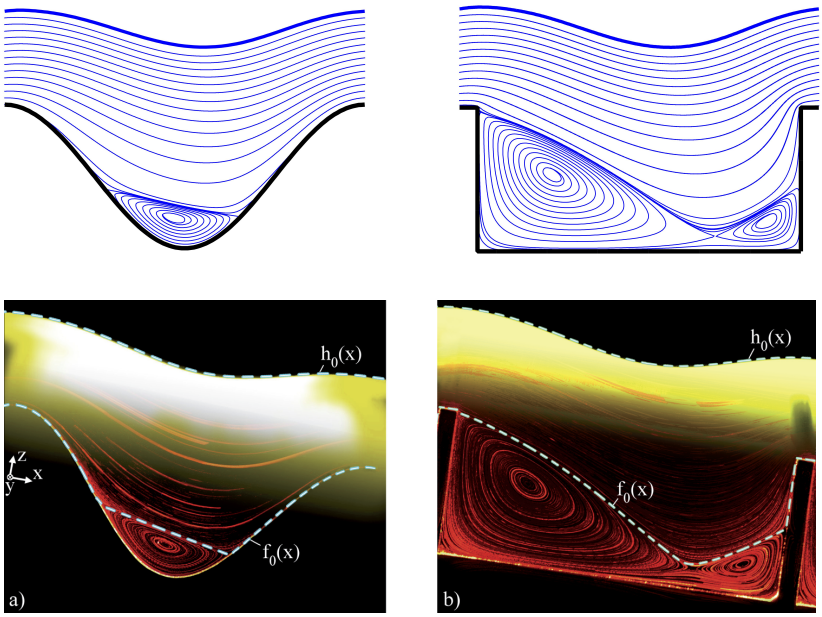


Figure 6.17: The upper streamline patterns reproduce the lower experimental results of [234] using silicon oil *Elbesil 145*: $\rho = 964.8\,\mathrm{kg/m^3}$, $\nu = 144.2\,\mathrm{mm^2/s}$, $\sigma = 20.1\,\mathrm{mN/m}$, $\lambda = 20\,\mathrm{mm}$, $A = 4\,\mathrm{mm}$, $H_0 = 5\,\mathrm{mm}$, $\alpha = 10°$; the corresponding non-dimensional numbers are $\mathrm{Re}_h = 5.12$, $\mathrm{Ca}_h = 1.022$.

substrate, various aspect ratios of the geometry and a wide range of inclination angles, Reynolds numbers and Capillary numbers. Although problems of mass conservation in conjunction with least-squares methods tend to increase when other than pure velocity boundary conditions are imposed, the example flows considered demonstrate the method to produce accurate results even in the case of periodic and free boundary conditions which pose a particular challenge.

Although a small number of selected parameter combinations is considered, some general trends can be observed. Initially a substrate contour of the form $b(x) = -a\cos(2\pi x/\ell)$ is considered. Fig. 6.19 shows, for fixed aspect ratios $\ell/h_0 = 5$, $a/h_0 = 1$ and non-dimensional numbers $\mathrm{Re}_h = 5$, $\mathrm{Ca}_h = 1$, the streamline patterns obtained for increasing inclination angle: $3° \leq \alpha \leq 45°$. As the inclination angle increases the free surface shape undergoes a complex transformation in which the shape minimum is shifted in the flow direction from the left to the right-hand side. For small angles, below 3°, the free surface is rather flat while its highest amplitude is reached when the free surface minimum passes the mid-point in the x-direction, $x = \ell/2$, which is approximately the case for $10° < \alpha < 12°$. When the inclination angle is increased further the free surface flattens again and for $\alpha > 20°$ both its shape and the internal flow structure remain essentially the same. In accord with the amplitude of the free surface shape the size of the eddy in the vicinity of

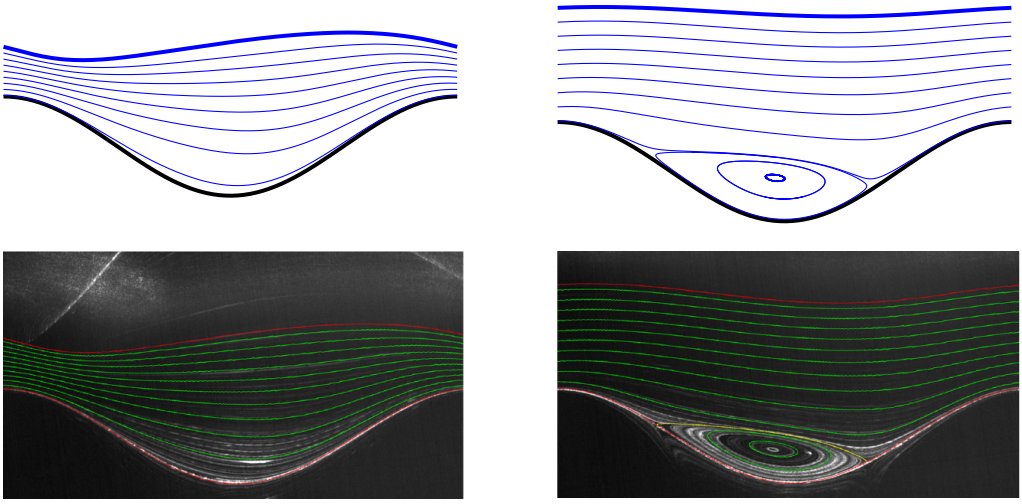


Figure 6.18: The upper streamline patterns reproduce the lower experimental results of Wierschem et al. [278] using silicon oil *Basildon-BC10* with the properties: $\rho = 924.3\,\mathrm{kg/m^3}$, $\nu = 11.6\,\mathrm{mm^2/s}$, $\sigma = 18.87\,\mathrm{mN/m}$. The geometry is given by $\lambda = 10\,\mathrm{mm}$, $A = 1\,\mathrm{mm}$, $\alpha = 8^\circ$ and $H_0 = 1\,\mathrm{mm}$ (left) / $2.2\,\mathrm{mm}$ (right). This translates to a Reynolds number $\mathrm{Re}_h = 5.66/55.67$ and a Capillary number $\mathrm{Ca}_h = 0.036/0.165$.

the substrate contour decreases in the range $3^\circ < \alpha < 12^\circ$ and increases again for $\alpha > 12^\circ$. Due to the fixed Reynolds number the tilt angle of the eddy is approximately constant.

Figure 6.20 shows, for $\ell/h_0 = 5$, $a/h_0 = 1$, $\alpha = 8^\circ$, $\mathrm{Ca}_h = 1$, the impact of varying Reynolds numbers ($2 \leq \mathrm{Re}_h \leq 50$) on the resulting flow structure. The transformation of the free surface shape is similar to the above case but in contrast the internal eddy grows with increasing Reynolds number and the eddy mid-point is shifted slowly in the flow direction; also the tilt angle of the eddy separatrix increases until approximately $\mathrm{Re}_h = 10$ and then decreases again together with a flattening of the free surface. Although thin gravity-driven film flow appears unstable at the free surface at a sufficiently high Reynolds number, as indicated in Haas [117], the region below the free surface in the vicinity of the eddy is heuristically stable and resolved correctly. These results are in accordance with the work of Haas [117], Pollak and Aksel [193], Scholle et al. [228].

Moreover, Fig. 6.21 shows, for $\ell/h_0 = 5$, $a/h_0 = 1$, $\alpha = 8^\circ$, $\mathrm{Re}_h = 5$, the impact of varying Capillary numbers $\mathrm{Ca}_h = 1/0.1/0.01$ on the resulting flow structure. The numerical results suggest that there is little change for Capillary numbers $\mathrm{Ca}_h > 1$ and $\mathrm{Ca}_h < 0.01$, that is most of the effect is in the range of investigation. The main effect of lowering the Capillary number is a smoothing and flattening of the free surface; for $\mathrm{Ca}_h = 0.01$ a nearly flat surface is reached which is accompanied by a moderate increase

of the eddy. Interestingly this also involves a slight shift of the free surface minimum in a direction opposite to that of the flow, i.e. to the left-hand side, which is opposite to the effect induced by increasing the Reynolds number or the inclination angle.

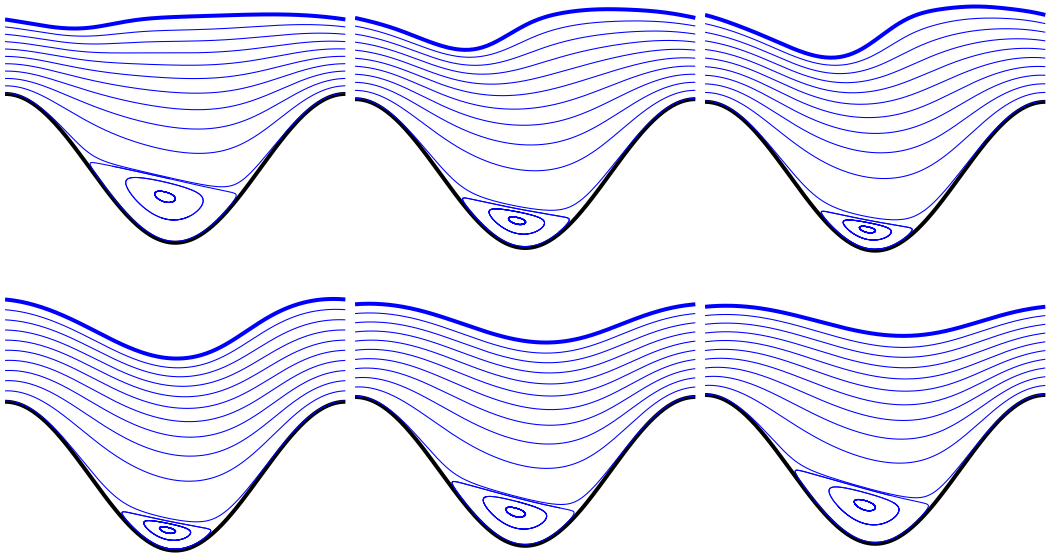


Figure 6.19: Film flow study with varying inclination angle: $\ell/h_0 = 5$, $a/h_0 = 1$, $\mathrm{Re}_h = 5$, $\mathrm{Ca}_h = 1$, $\alpha = 3°/8°/10°/12°/20°/45°$.

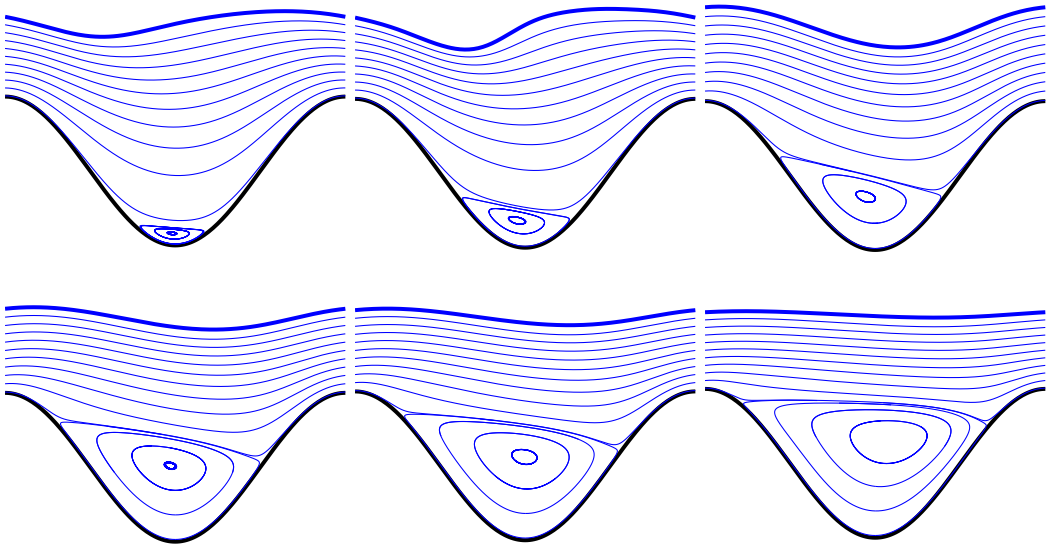


Figure 6.20: Film flow study with varying Reynolds number: $\ell/h_0 = 5$, $a/h_0 = 1$, $\alpha = 8°$, $\mathrm{Ca}_h = 1$, $\mathrm{Re}_h = 2/5/9/15/20/50$.

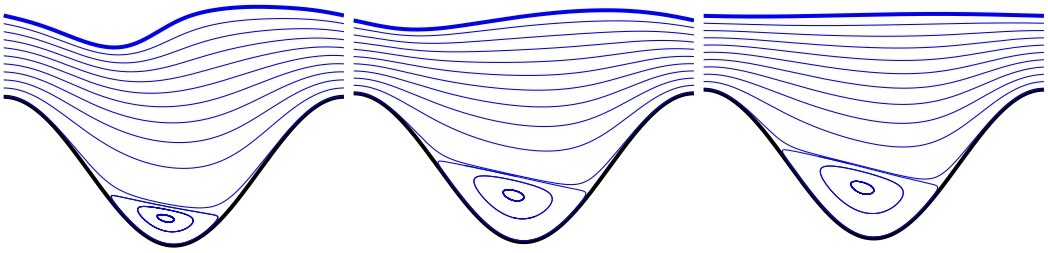


Figure 6.21: Film flow study with varying Capillary number: $\ell/h_0 = 5$, $a/h_0 = 1$, $\alpha = 8°$, $\mathrm{Re}_h = 5$, $\mathrm{Ca}_h = 1/0.1/0.01$.

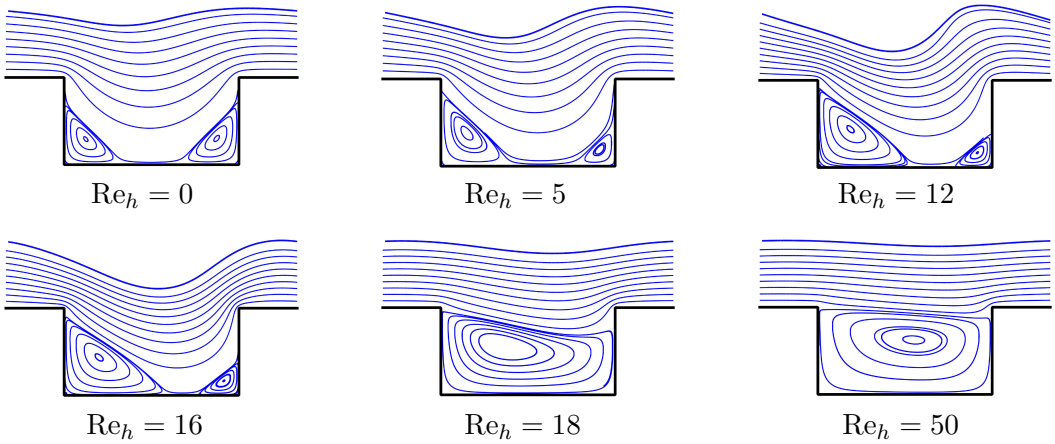


Figure 6.22: Thin gravity-driven film flow down an inclined substrate investigated in terms of varying the Reynolds number: $\ell/h_0 = 4$, $a/h_0 = 0.6$, $\alpha = 15°$ and $\mathrm{Ca} \to \infty$.

The FE code is in principle applicable for arbitrary substrate contour shapes. In Fig. 6.22 a study of film flow over a box-shaped contour at an angle of $\alpha = 15°$ degree and different Reynolds numbers is performed. At low Reynolds numbers the eddies in the corners are of equal size, with increasing Reynolds number the left one grows and the right one shrinks until they both coalesce to form one big eddy. The highest amplitude of the free surface occurs approximately at a Reynolds number of 12; afterwards the free surface becomes smoother. It is observed that those Reynolds numbers causing the strongest waviness of the surface (typically in the range of $5 < \mathrm{Re}_h < 20$ for the present application) also pose the most problems to the adaptive free surface trace algorithm, frequently requiring adjustment of parameters and generally a high number of iterations; this behaviour tends to get worse for decreasing Ca although one would expect that the corresponding flattening simplifies the finding of the free surface. On the other hand, stable Re and higher pseudo-stable Re generally lead to fast convergence.

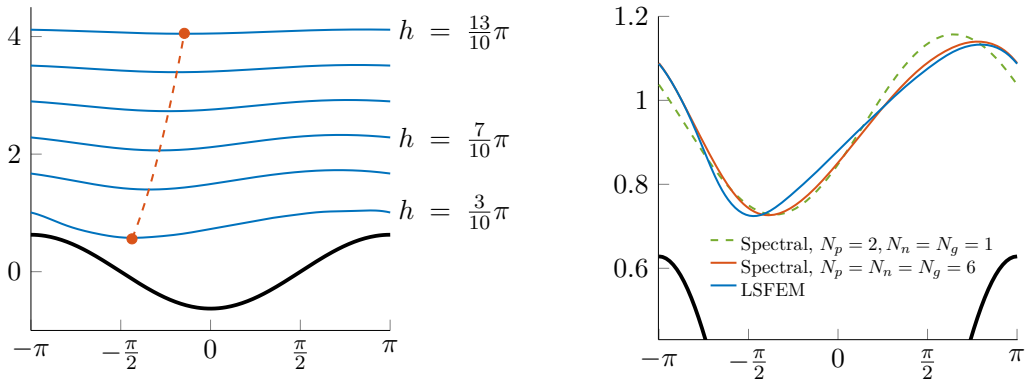


Figure 6.23: Impact of film thickness on free surface shape. Free surface profiles are computed by the spectral method, $N_p = N_n = N_g = 6$, for fixed $a = \pi/5$, $\alpha = 8°$, $\mathrm{Ca} = 1/\hat{\sigma} = 2$ and varying film height $h = (2k-1)\pi/10$, $k = 2, \ldots, 7$. The dotted red line connects the surface shape minima (left). Comparison between the spectral method, carried out for different mode numbers, and LSFEM for the above case $h = 3\pi/10$ (right); the axis are unequally scaled.

The same numerical approach has been used to explore the problem of continuous gravity-driven film flow down an inclined piece-wise planar and non-periodic substrate in the absence of inertia, Scholle et al. [232]. Numerical solutions of the first integral equations are compared with analytical ones from a linearised form of a reduced equation set resulting from application of the long-wave approximation. The results obtained are shown to: (i) be in very close agreement with existing, comparable experimental data and complementary numerical predictions for isolated step-like topography available in the open literature; (ii) exhibit the same qualitative behaviour for a range of Capillary numbers and step heights/depths, becoming quantitatively similar when both are small. A novel outcome of the formulation adopted is identification of an analytic criteria enabling a simple classification procedure for specifying the characteristic nature of the free surface disturbance formed; leading subsequently to the generation of a related, practically relevant, characteristic parameter map in terms of the substrate inclination angle and the Capillary number of the associated flow.

Now the spectral method is investigated which, in the implemented form, is confined to Stokes flow and sinusoidal substrate contours. Fig. 6.23 (left) demonstrates the impact of film thickness on the free surface contour. As the film thickness h is reduced[5], the amplitude of the free surface increases and the symmetry is increasingly broken; the latter leading to a shift of the minimum of the free surface contour in a direction opposite to the flow (to the left). A connection between the different shape minima (red dotted line)

[5]Here, again the wave length scaling without index h is used leading to $\mathrm{Ca} = \eta u_n/\sigma$ with u_n according to (6.12).

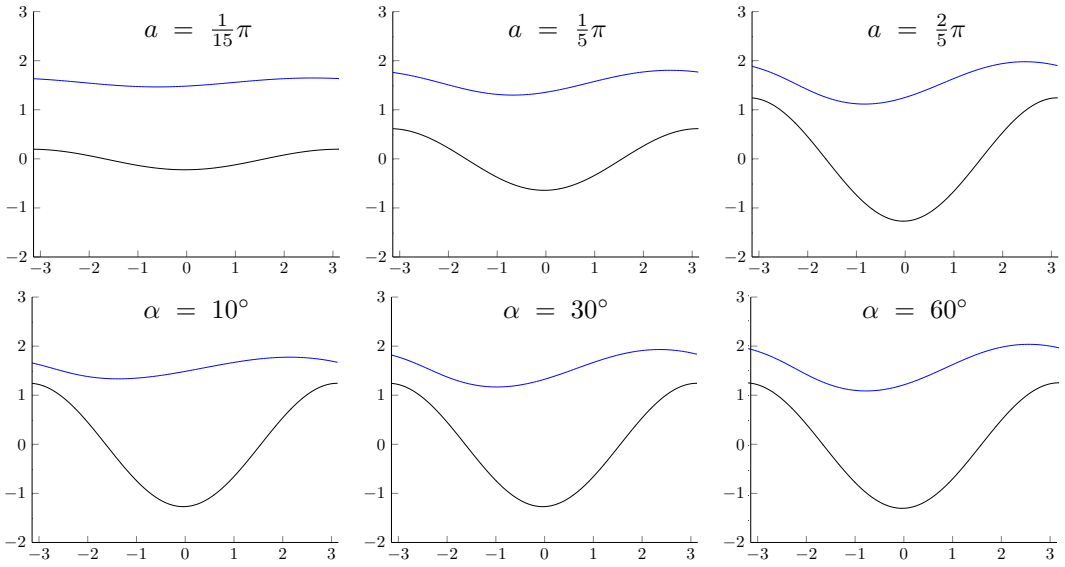


Figure 6.24: Film flow parameter study with fixed $h = \pi/2$, Ca $= 0.2$. Upper row: impact of substrate waviness on the free surface shape with $\alpha = 45°$ and $a = \pi/15$, $\pi/5$, $2\pi/5$. Bottom row: impact of the inclination angle on the free surface shape with $a = 2\pi/5$ and $\alpha = 10°$, $30°$, $60°$.

is well approximated by a simple quadratic function. While for a film height of $h = 1.3\pi$ a nearly flat free surface occurs, the amplitude for $h = 0.3\pi$ is already approximately 0.3, which is more than half of the substrate amplitude $a = \pi/5$. The results were calculated using a low number of modes, $N_p = N_n = N_g = 6$, and obtained significantly quicker as compared to the LSFEM; this is normally expected for a highly customised semi-analytic code.

However, while a comparison between the spectral method and the LSFEM leads to congruent results for near harmonic free surface shapes, which is normally the case for sufficiently high film thickness, very thin films lead to a significant deviation between both methods, as is shown for the case $h = 0.3\pi$ on the right-hand side of Fig. 6.23. The LSFEM result, obtained on a grid with approximately 50,000 nodes, can be considered very accurate and it is observed that the spectral method with a minimal number of modes, $N_p = 2$, $N_n = N_g = 1$, deviates strongly; when increasing the number of modes to $N_p = N_n = N_g = 6$, the deviation near the endpoints $\pm\pi$ is marginalised while near the free surface shape minimum, between $-\pi/2$ and 0, a severe discrepancy remains, especially with respect to the precise minimum location. The latter does not change when the number of modes is increased further to about 10. A problem of the spectral method is that a high number of modes (> 10) leads to badly conditioned matrices which are not solvable by standard 16 digit precision any more. Probably an increase of digits would

lead to a better result, but also the approximation (6.36) could be responsible for the persisting error. However, this is not investigated further here.

Finally some representative results obtained by the spectral method are shown. The upper row of Fig. 6.24 demonstrates the impact of the substrate waviness on the free surface shape. It is clearly seen that an increase in waviness also leads to an increase of the free surface amplitude and a shift of the surface minimum in the opposite direction to the flow (to the left). As expected, a flat substrate results in a flat, parallel free surface while, for example, a substrate amplitude of $a = 2\pi/5$ results in a free surface amplitude of 0.43 which is approximately 34% that of the substrate amplitude. The number of calculated positive and negative modes is $N_p = 15$ and $N_n = N_g = 6$. The bottom row of Fig. 6.24 shows the impact of the inclination angle on the free surface, calculated with the same number of modes as before.

6.3 Film flow over non-planar surfaces

Following the previous section on film flow over 2D corrugated surfaces it is important to highlight that the first integral methodology is also amenable to curvilinear surfaces as mentioned in Sec. 3.4.3. The advantages found for the 2D Cartesian case, i.e. that the equations of motion can be described in terms of two real-valued auxiliary potentials which allow for a convenient form of the dynamic boundary condition, extend to 3D domains with an arbitrary symmetry; for example axis-symmetric flows described in cylindrical or spherical coordinates. Although such flows are not considered in greater detail, an outlook is given below of how such symmetries can be treated within the first integral methodology in terms of the continuous coating of a hemisphere.

6.3.1 Steady axisymmetric flow over a hemisphere

When changing from Cartesian to spherical coordinates, the field equations (3.136), (3.137) and the dynamic boundary conditions (3.138), (3.139) have to be replaced by (3.142), (3.143) and (3.144), (3.145), respectively, as shown in Sec. 3.4.3. In principle, the least-squares FE methodology, both the theoretical part in Ch. 4 and the implementation part in Ch. 5 together with the multigrid solution strategy, Sec. 5.4, could be readily adapted for spherical coordinates to account for full inertial NS flow; however, this is beyond the scope of the thesis. Instead, the example of film flow over a hemisphere is considered in the context of the lubrication approximation only.

Consider a steady axisymmetric flow over a hemisphere of radius r_0, using spherical coordinates r, ϑ, φ with the respective velocities $u_1 = u_r$ and $u_2 = u_\vartheta$ given by (3.141), while $u_\varphi = 0$. The following asymptotic analysis makes use of assumptions that are usually

common for other lubrication-like problems:

(i) The flow direction is predominantly in the polar direction, hence $u_r \ll u_\vartheta$.

(ii) Changes of the flow in the polar direction are much smaller than changes in the radial direction, thus $\partial_\vartheta u_\mu \ll \partial_r u_\mu$, $\mu = 1, 2$.

(iii) For a thin film it can also be assumed $u_\mu / r \ll \partial_r u_\mu$.

(iv) Finally Stokes flow $\mathrm{Re} \ll 1$ is assumed, allowing quadratic terms to be neglected.

Applying the above simplifications to equation (3.143) gives, to leading order:

$$\partial_r \left[2\frac{\partial_\vartheta \Phi}{r} - \eta u_\vartheta \right] = 0 \,, \tag{6.49}$$

allowing for direct integration with respect to r, delivering:

$$\partial_\vartheta \Phi = \frac{\eta r}{2} u_\vartheta + r\Phi_1(\vartheta) \,, \tag{6.50}$$

with an unknown integration function $\Phi_1(\vartheta)$. Accordingly, $\partial_\vartheta \Phi$ can be substituted into (3.142), which to leading order yields:

$$\partial_r \left[\frac{1}{r} \left(\partial_r \Phi - \frac{3\eta}{2} u_r \right) \right] = \frac{\Phi_1'(\vartheta)}{r^2} \,,$$

which can again be integrated, leading to:

$$\partial_r \Phi = \frac{3\eta}{2} u_r - \Phi_1'(\vartheta) + r\Phi_2(\vartheta) \,, \tag{6.51}$$

and another integration function $\Phi_2(\vartheta)$. By computing the derivative of (6.50) with respect to r and the derivative of (6.51) with respect to ϑ, the potential Φ is eliminated by the difference $\partial_r(6.50) - \partial_\vartheta(6.51)$:

$$\frac{\eta}{2} \partial_r (r u_\vartheta) + \Phi_1(\vartheta) - \frac{3\eta}{2} \partial_\vartheta u_r + \Phi_1''(\vartheta) - r\Phi_2'(\vartheta) = 0 \,.$$

Neglecting again $\partial_\vartheta u_r$ in comparison to $\partial_r (r u_\vartheta)$, the equation can be integrated with respect to r, leading to:

$$\eta r u_\vartheta + 2r \left[\Phi_1''(\vartheta) + \Phi_1(\vartheta) \right] - r^2 \Phi_2'(\vartheta) = \Phi_3(\vartheta) \,,$$

giving a third integration function $\Phi_3(\vartheta)$, and implying the following analytical form for the velocity profile:

$$\eta u_\vartheta = r\Phi_2'(\vartheta) - 2 \left[\Phi_1''(\vartheta) + \Phi_1(\vartheta) \right] + \frac{1}{r} \Phi_3(\vartheta) \,, \tag{6.52}$$

as a combination of powers r^n with $n = -1, 0, 1$. Considering (3.141), the streamfunction is obtained by integrating:

$$\eta \frac{\partial \Psi}{\partial r} = \left[-r^2 \Phi_2'(\vartheta) + 2r \left[\Phi_1''(\vartheta) + \Phi_1(\vartheta) \right] - \Phi_3(\vartheta) \right] \sin\vartheta \,,$$

leading finally to:

$$\eta\Psi = \left[-\frac{r^3}{3} \Phi_2'(\vartheta) + r^2 \left[\Phi_1''(\vartheta) + \Phi_1(\vartheta) \right] - r\Phi_3(\vartheta) \right] \sin\vartheta + \Phi_4(\vartheta) \,, \tag{6.53}$$

and the integration function $\Phi_4(\vartheta)$. Thus an asymptotic solution of the field equations (3.142), (3.143) has been obtained with yet unknown functions $\Phi_1, \cdots, \Phi_4$ which have to be obtained from the boundary conditions: at the surface of the hemisphere, $r = r_0$, the no-penetration condition $\Psi(r_0, \vartheta) = 0$ and the no-slip condition $u_\varphi(r_0, \vartheta) = 0$ have to be considered, while at the free surface, $r = r_0 f(\vartheta)$, the kinematic boundary condition $\Psi\left(r_0 f(\vartheta), \vartheta\right) = -\dot{V}_0$ with volumetric flow rate $\dot{V}_0$ and the first integral of the dynamic boundary condition (3.144), (3.145) have to be fulfilled. Overall a set of five equations for five unknown functions, $\Phi_1, \cdots, \Phi_4$ and the shape of the free surface, $f(\vartheta)$, is obtained as provided in detail in Appendix D.2.1: by successive elimination of unknowns one ends up with two equations (D.10), (D.13) for two unknown functions Φ_2 and f.

It is useful to introduce the non-dimensional function:

$$F(\vartheta) := \frac{\Phi_2(\vartheta)}{\varrho g r_0} \,, \tag{6.54}$$

the Bond number Bo, the non-dimensional flow rate Q and the non-dimensional local film thickness as:

$$\mathrm{Bo} := \frac{\varrho g r_0^2}{\sigma} \,, \tag{6.55}$$

$$Q := \frac{2\eta \dot{V}_0}{\varrho g r_0^4} \,, \tag{6.56}$$

$$h(\vartheta) := f(\vartheta) - 1 \,, \tag{6.57}$$

respectively, by which the two remaining equations (D.10), (D.13) can be written in the non-dimensional form:

$$-\left(\frac{2}{3} + \frac{h}{2} \right) hF' - (1+h)h'F - \frac{h'}{2\mathrm{Bo}} + \frac{1}{2}(1+h)^2 h' \cos\vartheta = \frac{Q}{2h^2 \sin\vartheta} \,, \tag{6.58}$$

$$\frac{1 - h''}{2\mathrm{Bo}} - \frac{h^2}{2} F'' - hh' \left[2F' + \frac{1+h}{2} \sin\vartheta \right] + \left[1 + h - hh'' \right] \left[F - \frac{1+h}{2} \cos\vartheta \right] = 0 \,, \tag{6.59}$$

where $f'^2 = h'^2$ is again neglected. Having solved the ODEs (6.58), (6.59), the streamfunc-

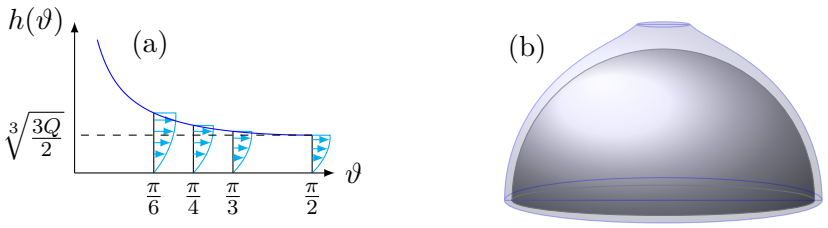


Figure 6.25: Film flow over a hemisphere: (a) film thickness vs. polar angle with velocity profile u_ϑ at fixed positions ϑ and (b) resulting film shape.

tion (D.3) and the polar velocity (D.4) are obtained as follows, when $\Phi_1'' + \Phi_1$ is eliminated according to (D.5) and the above scaling (6.54)–(6.57) is applied:

$$\frac{2\eta\Psi}{\varrho g r_0^4} = z^2 \left[\frac{2}{3}(h-z)F' \sin\vartheta - \frac{Q}{h^2}\right], \tag{6.60}$$

$$\frac{2\eta u_\vartheta}{\varrho g r_0^2} = \frac{2z}{1+z}\left[\left(z - \frac{2}{3}h\right)F' + \frac{Q}{h^2 \sin\vartheta}\right], \tag{6.61}$$

where:

$$z := \frac{r - r_0}{r_0},$$

has been introduced as a non-dimensional coordinate, indicating the distance from the surface of the hemisphere.

(a) Asymptotic solution for very thin films

Assuming that $h' \ll h \ll 1$, all the terms in (6.58), (6.59) containing h' are neglected and only the terms of leading order with respect to h are maintained, leading to the simplified set of equations:

$$-\frac{2}{3}hF' = \frac{Q}{2h^2 \sin\vartheta}, \tag{6.62}$$

$$\frac{1}{2\mathrm{Bo}} + F - \frac{1}{2}\cos\vartheta = 0, \tag{6.63}$$

with the obvious solution:

$$F(\vartheta) = \frac{1}{2}\cos\vartheta - \frac{1}{2\mathrm{Bo}}, \tag{6.64}$$

$$h(\vartheta) = \sqrt[3]{-\frac{3Q}{4F'(\vartheta)\sin\vartheta}} = \sqrt[3]{\frac{3Q}{2\sin^2\vartheta}}. \tag{6.65}$$

A visualisation of this solution is shown in Fig. 6.25: as expected the film becomes thinner for increasing polar angle and the appearance of the velocity profiles is close to the

parabolic profile known from the Nusselt solution for film flow over an inclined plane.

It is remarked that the result for the film thickness (6.65) corresponds to that of [246], as demonstrated in Appendix D.2.2. On the other hand the velocity profile (6.61) is not exactly parabolic as assumed by Takagi and Huppert [246] and not even self-similar, i.e. u_ϑ can not be written in a form $h^n f(z/h)$.

6.4 Unsteady Couette flow confined between two corrugated rigid surfaces

The case of Couette flow for a Newtonian, incompressible fluid confined between two horizontally aligned rigid surfaces, both corrugated sinusoidally and driven by the movement of the upper surface with uniform velocity U_0, the lower one remaining stationary, is explored, see Marner et al. [172]. The flow configuration is illustrated schematically in Fig. 6.26; the unsteady character of the flow being due to the geometry of the domain varying with time. The geometry is defined in terms of the mean film thickness H_0, the amplitudes A and H_1 of the lower and upper surfaces, respectively, and the wavelength λ for both surfaces. External forces are not considered subsequently.

Exploration of the above unsteady Couette problem ties in well with an existing series of experimental and numerical investigations on thin film and channel flows over periodically occurring topography [170, 228, 234, 258], which is a topic of considerable interest in diverse technical and industrial processes; for example, in coating [147] or lubrication [170, 217, 229] applications, heat exchangers and evaporators [245]. The existence of isolated or periodically occurring topographical features as exemplified by the sinusoidal lower surface contour in Fig. 6.26, can give rise to the formation of closed eddy structures [105, 239, 267, 276] leading to particle trapping and stagnant flow in separated flow regions. In the subsequent investigation the reader's attention is directed in particular to the inner flow structure present within the valleys of the lower surface topography and to the mechanism of mass exchange from the fluid in such valleys to the overlying bulk flow and vice versa, as the channel thickness varies with time.

The effect of mass exchange together with the prospects of process enhancement have been studied by several authors: for instance, Wierschem and Aksel [276] observed experimentally the transport of inert tracers from fluid in the valleys of sinusoidal topography in the presence of surface waves, enabled via a turnstile lobe mechanism; Horner et al. [129] provides a comprehensive overview of this mechanism for modulated flow over a square cavity, while Wilson et al. [279] investigated, both experimentally and theoretically, the enhancement of transport and stirring between two rollers via lobe dynamics. Although these research topics are in the main associated with flows involving a free surface, the

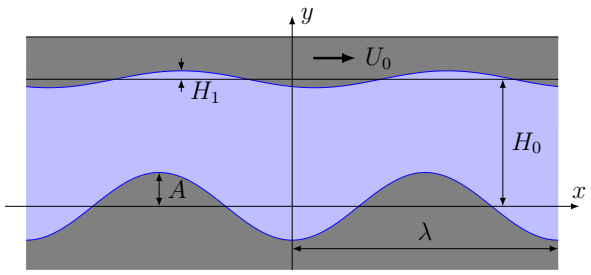


Figure 6.26: Schematic of the unsteady Couette flow configuration, showing the relevant defining geometrical parameters.

turnstile lobe effect is purely driven by temporal changes of the geometry of the flow domain, which can conveniently and more easily be studied in the model framework of Couette flow, in which the geometry variance is artificially induced by specifying a non-uniform moving upper surface; that is, one with a well defined topography profile as shown in Fig. 6.26.

The boundary value problem as defined is formulated in (Sec. 6.4.1) and the effect of mass exchange within the channel explored by means of two different self-contained methods, highlighting in addition different ways in which the new reformulation of the NS equations might be beneficially utilised. The first (Sec. 6.4.2) is based on a generalised form of the Goursat representation (1.33), combined with a spectral Fourier discretization of the holomorphic functions involved; the second (Sec. 6.4.3) is a FE procedure based on a weak integral form of the field equations. The corresponding results are presented in Sec. 6.4.4.

6.4.1 Equations of motion

All relevant quantities are scaled in terms of $\lambda/(2\pi)$, U_0, $\lambda/(2\pi U_0)$ and $2\pi\eta U_0/\lambda$ for lengths, velocities, the time and for the pressure, respectively. Consequently, the set of relevant parameters is reduced to two non-dimensional amplitudes a, h_1, a non-dimensional film thickness h_0 and the Reynolds number Re given by:

$$\mathrm{Re} := \frac{\varrho U_0 \lambda}{2\pi\eta}\,, \qquad a := \frac{2\pi A}{\lambda}\,, \qquad h_0 := \frac{2\pi H_0}{\lambda}\,, \qquad h_1 := \frac{2\pi H_1}{\lambda}\,;$$

accordingly, the scaled governing field equations (3.12), (3.13) read:

$$-\mathrm{Re}\frac{u^2}{2} \;=\; 4\frac{\partial^2\chi}{\partial\bar{\xi}^2}\,, \tag{6.66a}$$

$$-\mathrm{i}\left[\mathrm{Re}\frac{\partial\Psi}{\partial t} - 4\frac{\partial^2\Psi}{\partial\bar{\xi}\partial\xi}\right] + p + \mathrm{Re}\frac{\bar{u}u}{2} \;=\; 4\frac{\partial^2\chi}{\partial\bar{\xi}\partial\xi}\,. \tag{6.66b}$$

The two boundaries formed by the lower $\xi = \beta(x)$ and upper $\xi = \varphi(x,t)$ corrugated surfaces are given by the functions:

$$\beta(x) := x - \mathrm{i}a\cos x\,, \tag{6.67}$$

$$\varphi(x,t) := x + \mathrm{i}h_0 - \mathrm{i}h_1\cos(x-t)\,, \tag{6.68}$$

along which the following no-slip/no-penetration conditions have to be fulfilled:

$$u\left(\beta(x),\bar{\beta}(x),t\right) = 0\,, \tag{6.69}$$

$$u\left(\varphi(x,t),\bar{\varphi}(x,t),t\right) = 1\,. \tag{6.70}$$

As an initial condition the fluid is assumed to be at rest, that is:

$$\Psi\left(\xi,\bar{\xi},t_0\right) = \chi\left(\xi,\bar{\xi},t_0\right) = 0\,. \tag{6.71}$$

With reference to the work of Scholle et al. [229] and Esquivelzeta-Rabell et al. [91], who have shown for steady Couette flow generated with a flat upper driving surface that inertial effects play a minor role only up to Reynolds numbers with a value of about 10, by restricting the current investigation to Reynolds number $\mathrm{Re} \leq 10$ the nonlinear terms in the governing field equations (6.66a), (6.66b) can be effectively omitted while retaining the Reynolds number in the term involving the time derivative, leading to:

$$p - \mathrm{i}\left[\mathrm{Re}\frac{\partial\Psi}{\partial t} - 4\frac{\partial^2\Psi}{\partial\bar{\xi}\partial\xi}\right] = 4\frac{\partial^2\chi}{\partial\bar{\xi}\partial\xi}\,, \tag{6.72a}$$

$$\frac{\partial^2\chi}{\partial\bar{\xi}^2} = 0\,. \tag{6.72b}$$

Below, variants of time-dependent Couette flow, involving aspect ratios of $a < h_0 < 2\pi$ and $h_1 \ll a$, are considered.

6.4.2 Asymptotic model and method of solution

(a) Generalised Goursat form

According to earlier studies [217, 221], the streamfunction for the case of steady Stokes flow with $h_1 = 0$ can conveniently be written as:

$$\Psi_\mathrm{s} = By^2 + 2\mathrm{Re}\left[R(\xi) + yQ(\xi)\right]\,, \tag{6.73}$$

where the constant B and the two 2π-periodic holomorphic functions R and Q are determined by the boundary conditions. By identifying $g_0(\xi) = 2R(\xi) - \mathrm{i}\xi Q(\xi) - B\xi^2/2$ and $g_1(\xi) = \mathrm{i}Q(\xi) + B\xi/2$, it becomes obvious that (6.73) is a variant of the Goursat form

(1.33), by which it is guaranteed that Ψ_s fulfils the biharmonic equation, $\Delta^2\Psi_\mathrm{s}=0$.

The streamfunction Ψ for unsteady flow with $h_1>0$ is expected to differs from (6.73). However, neglecting inertia and assuming small Reynolds numbers, $\mathrm{Re}<1$, it is proven, see Appendix D.3.1, that $\Delta^3\Psi=\mathcal{O}\left(\mathrm{Re}^2\right)$, motivating the following analytical 'triharmonic' form for Ψ:

$$\Psi = By^2 + 2\mathrm{Re}\left[R(\xi,t) + yQ(\xi,t) + \mathrm{Re}\frac{y^2}{2}P(\xi,t)\right]\,, \tag{6.74}$$

which is obviously a generalisation of (6.73) containing a third 2π-periodic holomorphic function P and considering time-dependence for the three functions P,Q,R and the parameter B. The analytical form $\partial\chi/\partial\bar{\xi}=f(\xi,t)$ is obtained directly from (6.72b). Inserting this and (6.74) into the field equation (6.72a) and neglecting terms of order $\mathcal{O}\left(\mathrm{Re}^2\right)$, the following identity is implied for the pressure field:

$$p = 4f' - 2\mathrm{i}B - \mathrm{iIm}\left\{-4Q' - \mathrm{Re}\left[2\mathrm{i}\left(\dot{R}-P\right) + 2y\left(\mathrm{i}\dot{Q}+2P'\right)\right]\right\}\,, \tag{6.75}$$

where the prime denotes a derivative with respect to ξ. Since the pressure has to be real-valued, the first three terms $4f'-2\mathrm{i}B-\mathrm{i}\dot{B}y^2$ have to equate to the complex expression inside the curly parentheses, leading to:

$$4f' = 2\mathrm{i}B - 4Q' - \mathrm{Re}\left[2\mathrm{i}\left(\dot{R}-P\right) + 2y\left(\mathrm{i}\dot{Q}+2P'\right)\right] + p_0(t)\,. \tag{6.76}$$

By taking the derivative of the above equation with respect to $\bar{\xi}$ and considering $\partial y/\partial\bar{\xi}=\mathrm{i}/2$, the identity:

$$\mathrm{i}\dot{Q} + 2P' = 0\,, \tag{6.77}$$

is obtained, revealing that the two complex functions P and Q take the form of a PDE.

(b) Perturbation approach and boundary conditions

According to the aforementioned restriction $h_1\ll 1$, the decomposition:

$$\begin{aligned} Q(\xi,t) &= Q_\mathrm{s}(\xi) + h_1 Q_\mathrm{u}(\xi,t)\,,\\ R(\xi,t) &= R_\mathrm{s}(\xi) + h_1 R_\mathrm{u}(\xi,t)\,,\\ P(\xi,t) &= h_1 P_\mathrm{u}(\xi,t)\,, \end{aligned}$$

of the three complex functions P,Q,R is applied, where the subscript 's' denotes the corresponding steady flow ($h_1=0$) and 'u' the small perturbation invoked by the moving, slightly corrugated, upper surface. As a consequence, the complex conjugate of the complex velocity field is likewise decomposed as $\bar{u}=2\mathrm{i}\partial\Psi/\partial\xi=\bar{u}_\mathrm{s}+h_1\bar{u}_\mathrm{u}$ with:

$$\bar{u}_\mathrm{s} := 2By + 2\mathrm{i}\left[R_\mathrm{s}{}' + yQ_\mathrm{s}{}'\right] + 2\mathrm{Re}\,Q_\mathrm{s}\,, \tag{6.78}$$

$$\bar{u}_\mathrm{u} := 2\mathrm{i}\left[{R_\mathrm{u}}' + y{Q_\mathrm{u}}'\right] + \mathrm{Re}\frac{y^2}{2}\dot{Q}_\mathrm{u} + 2\mathrm{Re}\left[Q_\mathrm{u} + \mathrm{Re}\, yP_\mathrm{u}\right] . \tag{6.79}$$

Considering the above, the complex conjugate of the no-slip/no-penetration condition (6.69) at the lower surface can be decomposed into the two boundary conditions:

$$\bar{u}_\mathrm{s}\left(\beta(x), \bar{\beta}(x)\right) = 0\,, \tag{6.80}$$

$$\bar{u}_\mathrm{u}\left(\beta(x), \bar{\beta}(x), t\right) = 0\,; \tag{6.81}$$

while at the upper surface a domain perturbation is given by (6.68) due to h_1 appearing explicitly. Via a Taylor expansion of the no-slip/no-penetration condition (6.70) with respect to h_1 and sorting terms by powers of h_1, one ends up with the two conditions:

$$\bar{u}_\mathrm{s}\left(x + \mathrm{i}h_0, x - \mathrm{i}h_0\right) = 1\,, \tag{6.82}$$

$$\bar{u}_\mathrm{u}\left(x + \mathrm{i}h_0, x - \mathrm{i}h_0, t\right) = \mathrm{i}\left[\frac{\partial \bar{u}_\mathrm{s}}{\partial \xi} - \frac{\partial \bar{u}_\mathrm{s}}{\partial \bar{\xi}}\right]\cos(x - t)\,. \tag{6.83}$$

With reference to Appendix D.3.2, these two equations reveal a hierarchy: the inhomogeneity in equation (6.83) for the first order perturbation depends on the base solution. Apart from this, the inhomogeneity is purely harmonic with respect to time. Due to the linearity of the problem, the perturbation must be harmonic with respect to time too, implying the following analytical form:

$$\begin{aligned} P_\mathrm{u}(\xi, t) &= p^+(\xi)\exp(\mathrm{i}t) + p^-(\xi)\exp(-\mathrm{i}t)\,, \\ Q_\mathrm{u}(\xi, t) &= q^+(\xi)\exp(\mathrm{i}t) + q^-(\xi)\exp(-\mathrm{i}t)\,, \\ R_\mathrm{u}(\xi, t) &= r^+(\xi)\exp(\mathrm{i}t) + r^-(\xi)\exp(-\mathrm{i}t)\,, \end{aligned}$$

for the three holomorphic functions. By inserting the above forms into equation (6.77), it follows that:

$$2{p^\pm}'(\xi) = \pm q^\pm(\xi)\,. \tag{6.84}$$

(c) Discretisation by Fourier decomposition

In line with the work of others [170, 217], the periodic holomorphic functions are represented by a truncated Fourier series as follows:

$$Q_\mathrm{s}(\xi) = \sum_{k=-N}^{N} Q_k \exp(\mathrm{i}k\xi)\,, \qquad q^\pm(\xi) = \sum_{k=-N}^{N} q_k^\pm \exp(\mathrm{i}k\xi)\,, \tag{6.85}$$

$$R_\mathrm{s}(\xi) = \sum_{k=-N}^{N} R_k \exp(\mathrm{i}k\xi)\,, \qquad r^\pm(\xi) = \sum_{k=-N}^{N} r_k^\pm \exp(\mathrm{i}k\xi)\,, \tag{6.86}$$

up to order $N \in \mathbb{N}$, reducing the problem to a finite set of complex coefficients and the yet unknown constant B. For the two remaining functions $p^{\pm}(\xi)$, equations (6.84) are fulfilled identically by:

$$p^{\pm}(\xi) = p_0 \mp \mathrm{i} \sum_{\substack{k=-N \\ k\neq 0}}^{N} \frac{q_k^{\pm}}{2k} \exp(\mathrm{i}k\xi)\,, \tag{6.87}$$

with integration constant p_0. On inserting the above series representation (6.85)-(6.87) into boundary conditions (6.80), (6.81) and (6.82), (6.83), a linear set of algebraic equations for the coefficients $Q_k, R_k, q_k^{\pm}, r_k^{\pm}, p_0$ and B is obtained. Full details regarding the formulation of this algebraic set of equations is provided in Appendix D.3.2.

6.4.3 Numerical model and method of solution

The starting point for a weak integral formulation is again the linear field equations (6.72a), (6.72b). As the pressure is not of relevance for the problem under investigation, the imaginary part only of (6.72a) is taken into account, namely:

$$\frac{\mathrm{Re}}{4}\frac{\partial \Psi}{\partial t} - \frac{\partial^2 \Psi}{\partial \bar{\xi}\partial \xi} = -\,\mathrm{Im}\left(\frac{\partial^2 \chi}{\partial \bar{\xi}\partial \xi}\right) =: -\Phi(\xi, \bar{\xi}, t)\,, \tag{6.88}$$

and from (6.72b) it follows directly that:

$$\mathrm{Im}\left(\frac{\partial^4 \chi}{\partial \bar{\xi}^2 \partial \xi^2}\right) = \frac{\partial^2 \Phi}{\partial \bar{\xi}\partial \xi} = 0\,, \tag{6.89}$$

in which the more convenient real-valued field Φ has been introduced, thus replacing the complex-valued χ.

In order to solve for the modified field equations (6.88), (6.89), complemented by the conditions (6.69)-(6.71), the implicit Crank-Nicolson time discretization scheme is combined with a weak Galerkin FE formulation, resulting in a second order accurate method in both space and time. Consequently, iteration in time, starting with $\Psi_0 = \Phi_0 = 0$, is accomplished according to:

$$\frac{\mathrm{Re}}{2\Delta t}\Psi_{t+1} - \frac{\partial^2 \Psi_{t+1}}{\partial \bar{\xi}\partial \xi} + \Phi_{t+1} = \frac{\mathrm{Re}}{2\Delta t}\Psi_t + \frac{\partial^2 \Psi_t}{\partial \bar{\xi}\partial \xi} - \Phi_t\,, \tag{6.90}$$

$$\frac{\partial^2 \Phi_{t+1}}{\partial \bar{\xi}\partial \xi} = 0\,, \tag{6.91}$$

and at each time step a spatial FE problem is solved based on the weak variational formulation:

Find $\Psi \in \{\Psi \in H^1(\Omega) : \Psi = g_1 \text{ on } \partial\Omega\}$ and $\Phi \in H^1(\Omega)$,
such that for all $v \in H^1(\Omega)$ and $w \in H_0^1(\Omega)$:

$$\frac{\mathrm{Re}}{2\Delta t}\langle \Psi_{t+1}, v\rangle_{L_2,\Omega} - \langle \nabla\Psi_{t+1}, \nabla v\rangle_{L_2,\Omega} + \langle \Phi_{t+1}, v\rangle_{L_2,\Omega} = \langle g_2, v\rangle_{L_2,\partial\Omega} + b_{t+1}(v, \Psi_t)\,, \quad (6.92)$$

$$b_0 = 0\,, \quad b_{t+1}(v, \Psi_t) = \frac{\mathrm{Re}}{\Delta t}\langle \Psi_t, v\rangle_{L_2,\Omega} - b_t\,, \quad (6.93)$$

$$\langle \nabla\Phi_{t+1}, \nabla w\rangle_{L_2,\Omega} = 0\,, \quad (6.94)$$

in which the standard L_2-inner product is used and the standard H^1 Sobolev space as defined in Appendix B.1.2. H_0^1 comprises H^1-functions with zero boundary conditions. Ω and $\partial\Omega$ denote the computational domain and its boundary, respectively; the functions g_1 and g_2 define the space and time-dependent velocity boundary conditions for the streamfunction.

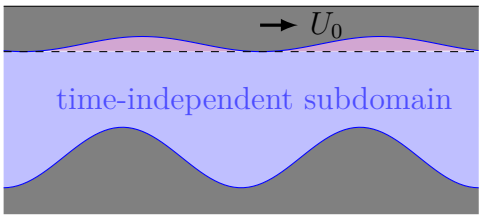


Figure 6.27: Restriction of the problem to a time-independent domain.

In order to provide a complete description of the solution procedure it is necessary to take a closer look at the boundary conditions. First of all, the no-slip and no-penetration conditions at the lower, stationary surface lead to $g_1(x, \beta(x), t) = g_2(x, \beta(x), t) = 0$, in which the arbitrary constant offset of the streamfunction is fixed. The periodicity of the domain is incorporated simply by identifying the related nodes on the left and right-hand side of the FE grid; therefore a separate periodic boundary treatment can be omitted. In order to avoid complications with the upper moving bounding surface and any time-dependent re-meshing issues, condition (6.68), (6.70) is approximated as a flat surface with a constant hight of $h_0 - h_1$, that is directly *beneath* the surface waviness. In this sense the impact of the lower surface corrugations on the flow near the upper surface is neglected so that for the approximation of Ψ close to the height $h_0 - h_1$, a geometry with $\tilde{h}_0 = h_0$, $\tilde{h}_1 = h_1$ and $\tilde{a} = 0$ is assumed, see Fig. 6.27.

An analytical approximation of the above reference problem can be derived under the assumptions of lubrication theory and for moderate aspect ratios [217], giving:

$$\Psi(x, h_0 - h_1, t) \approx U_0 y - \Psi_0(x - t, h_0 - h_1) + c_0\,, \quad (6.95)$$

$$\Psi_0(x, y) = \frac{h_0(h_0^2 - h_1^2)}{2h_0^2 + h_1^2} Y^2(3 - 2Y) - h_0 Y^2(1 - Y)\left(1 + \frac{h_1}{h_0}\cos(x)\right)\,,$$

$$Y(x, y) := \frac{y + h_1\cos(x)}{h_0 + h_1\cos(x)}\,.$$

The boundary functions g_1 and g_2 can than be derived from (6.95) and its normal derivative. In order to fix the remaining constant c_0, note that an additional constraint for the vorticity ω is assigned to the upper surface as is similarly the case for all multiply connected domains [250]. Denoting the upper boundary as Γ_u and the tangential and normal velocity components with subscripts τ and n, respectively, then by integrating the equation:

$$\mathrm{Re}\left[\frac{\partial u_\tau}{\partial t}+u_\tau\frac{\partial u_\tau}{\partial \tau}\right]=-\frac{\partial p}{\partial \tau}+\frac{\partial \omega}{\partial n}\,,$$

over Γ_u, the condition:

$$-\frac{1}{\mathrm{Re}}\int_{\Gamma_u}\frac{\partial \omega}{\partial n}\,\mathrm{d}\Gamma_u=\frac{\partial u_\tau}{\partial t}\int_{\Gamma_u}\mathrm{d}\Gamma_u+\int_{\Gamma_u}\frac{\partial}{\partial \tau}\left[\frac{1}{2}u_\tau^2+p\right]\,\mathrm{d}\Gamma_u=0\,,$$

is established, which in the present context translates as the following condition on the potential Φ:

$$0=\int_{\Gamma_u}\frac{\partial \omega}{\partial n}\,\mathrm{d}\Gamma_u=\int_{\Gamma_u}\frac{\partial}{\partial n}\left[\frac{\mathrm{Re}}{4}\frac{\partial \Psi}{\partial t}+\Phi\right]\,\mathrm{d}\Gamma_u=\int_{\Gamma_u}\frac{\partial \Phi}{\partial n}\,\mathrm{d}\Gamma_u\,.$$

6.4.4 Results and discussion

(a) Asymptotic Results

According to Appendix D.3.2(a) the original set of $4N+3$ algebraic equations is reduced to the set (D.24) of $2N+1$ equations for the coefficients $Q_n, R_n; n>0$ and Q_0. This set of equations was solved for $a=\pi/8$ and $h_0=\pi/4$ using Maple at a truncation order of $N=24$. The remaining coefficients are determined according to (D.21), (D.22) and (D.20). A plot of the resulting streamline pattern is shown in Fig. 6.28: note that the presence of a closed eddy in the valley formed by the lower surface topography is in complete accordance with prior results reported in literature [170, 217, 221, 229].

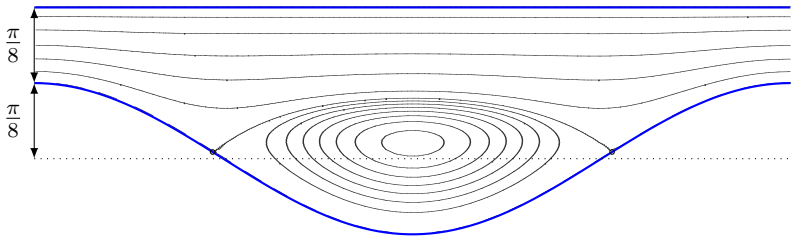


Figure 6.28: Steady Couette flow ($h_1=0$) for $a=\pi/8$ and $h_0=\pi/4$, Re $=0$: streamlines revealing the associated flow structure. Stationary points are indicated by a bold dot.

The coefficients $q_k^{\pm}$ and $r_k^{\pm}$, by which the complex functions for the first perturbation order are determined, result from solving the set of equations (D.25) and (D.27) given in Appendix D.3.2(b). Again, Maple is used to solve the associated algebraic equation set. Since the perturbation provides a small contribution only to the entire solution, a truncation order of $N = 12$ proves sufficiently accurate.

As for the steady Stokes flow case, the same geometric parameters are chosen for $a = \pi/8$ and $h_0 = \pi/4$; the additional parameter for unsteady flow, being the choice $\mathrm{Re} = 0.5$ and thus $h_1 = 0.02$. The resulting instantaneous streamline patterns at four different times,

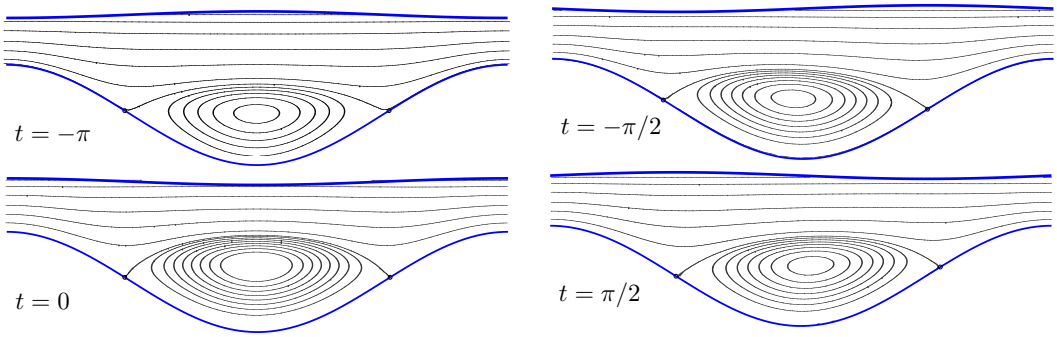


Figure 6.29: Unsteady Couette flow with $a = \pi/8$, $h_0 = \pi/4$, $\mathrm{Re} = 0.5$ and $h_1 = 0.02$; instantaneous streamline patterns, at four different times, revealing the associated flow structure.

$t \in \{-\pi, -\pi/2, 0, \pi/2\}$, are shown in Fig. 6.29: the time-dependence of the flow invoked by the corrugated upper surface is considerable, especially regarding the shape and skewness of the eddy, regardless of the small amplitude h_1. Since during each time period the eddy occupies slightly different regions of the flow domain, it is only to be expected that some fluid particles which are located in the vicinity of the border between the eddy and main flow, while trapped inside (outside) of the eddy at some point in time will be located outside (inside) the eddy at some different point in time, due to an associated shift of the separating streamline, or separatrix. Accordingly, the mechanism for material exchange between an eddy and the main flow is captured qualitatively, see Fig. 6.30. In order to study this feature in more detail, the movement of material particles has to be visualised by path lines or sweep lines. This is conveniently performed numerically as described below.

(b) Numerical results

FE calculations based on the weak formulation (6.92)-(6.94) were performed within a Matlab framework, with a triangular mesh structure containing of the order of 40 thousand elements and employing quadratic Lagrange elements for all test and solution spaces; one time period, $T = \lambda/U_0$, is discretized over 150 time steps.

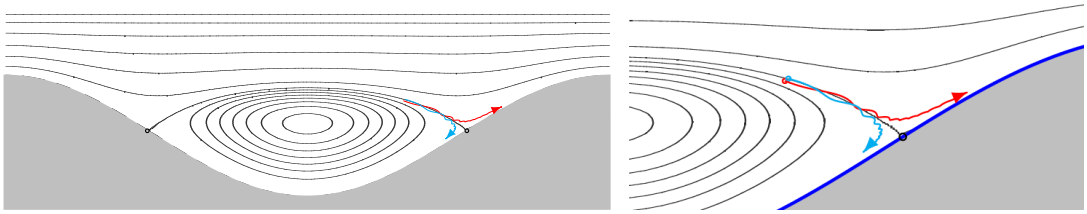


Figure 6.30: For the same geometry and conditions as in Fig. 6.29, the movement of two single particles is tracked for several 2π-time periods; the red particle leaves the eddy and is flushed away with the bulk flow while the blue one is approaching from the bulk flow and then entrapped in the eddy.

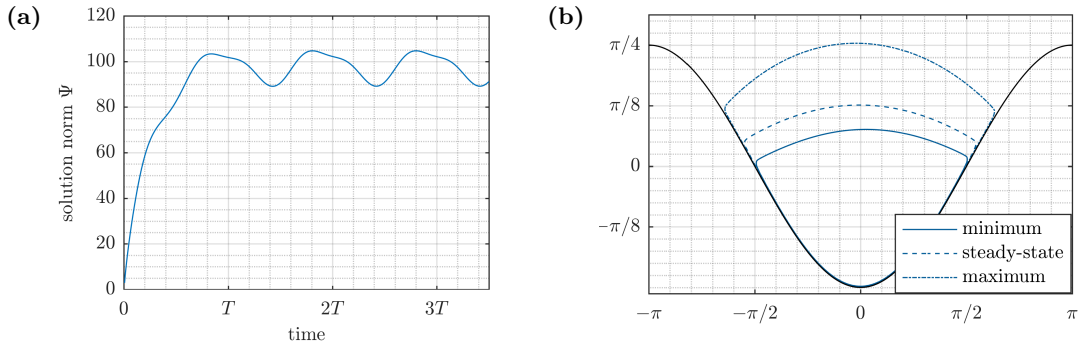


Figure 6.31: Couette flow between corrugated surfaces for the case of $\lambda = 2\pi$, $a = \pi/4$, $h_0 = \pi$ and $h_1 = a/20$ with a Reynolds number $\mathrm{Re} = 10$. (a) Euclidean norm of the FE solution vector of Ψ evolving with time; a time-periodic flow is established after a $2T$ time period. (b) Separatrices confining the eddy region from the bulk flow for three different states; these being minimum and maximum eddy shape in terms of surface area occurring in the time-periodic regime as well as the steady-state eddy shape corresponding to the case of $h_1 = 0$.

First a representative example geometric configuration with $\lambda = 2\pi$, $a = \pi/4$, $h_0 = \pi$ and $h_1 = a/20$, for a Reynolds number of $\mathrm{Re} = 10$, is considered. Figure 6.31a shows that a time periodic flow field is established after just two time periods and that in this time-periodic regime the size and shape of the eddy lies either side the equivalent steady state configuration which exists when $h_1 = 0$, see Fig. 6.31b. In this context the question arises whether particles entrapped within the steady-state eddy can be made to escape by being flushed away by the bulk flow. The subsequent investigations indicate that in general a mass exchange, at least in both directions, cannot be expected for arbitrarily small amplitudes h_1 of the upper plate. To visualise material transport the time evolution of sweep lines was computed, i.e. material lines consisting of the same fluid particles at all times, their initial shape being conveniently defined by corresponding streamlines at a particular time.

Figure 6.32 shows results for the above geometry in more detail, in the form of a sequence of snap-shots. At the time $t = 0$ the corrugated upper surface is considered to instantaneously accelerate to a velocity of U_0 and then the flow is subsequently solved for over 20 time periods; the direction of flow is always from left to right. At $t = 3.44\,T$, representing a state with a maximum eddy size, the instantaneous streamline pattern is captured in terms of the 'initial' sweep lines present, the *material movement* of which is then tracked in the following sequence of plots; for each sweep line the advection equation $\frac{\mathrm{d}x}{\mathrm{d}t} = u(x,t)$ is solved for the trajectories $x(t)$ via a fourth-order Runge-Kutta scheme. Each sweep line is represented by between 500 and 1000 fluid particles. Moreover, the shaded area represents the material time evolution of the fluid region which at time $t = 3.44\,T$ is initially defined by the *shape of the original closed steady-state eddy*. From the deformation of the sweep lines it can be seen that material from the maximum eddy is flushed away with the bulk flow and fluid in the region of its bounding separatrix plunges lower down the right hand side of the valley to displace the separatrix bounding the shaded former steady-state area. Part of the fluid originally in the maximum eddy is entrained into the shaded region; however, over the course of solution the material within the steady state eddy region is not expelled and remains trapped. These observations clearly indicate that small perturbations of the steady-state flow, interpreted as being due to the corrugations in the upper plate, leave the material movement of the steady-state flow qualitatively invariant: a certain amount of fluid particles covering an area smaller than the maximum eddy size but larger than the minimum eddy size, as expected fairly close to the shape of the steady-state eddy, remain trapped; whereas fluid particles above this region are flushed away.

In contrast to the above, the following example considers the effect of varying the amplitude of the upper plate h_1. Figures 6.33-6.35 show sweep lines for the configuration $\lambda = 2\pi$, $a = \pi/8$, $h_0 = \pi/4$, $\mathrm{Re} = 1$ with different amplitudes $h_1 = 0.01$, $h_1 = 0.03$ and $h_1 = 0.09$. As before, several time periods are allowed to pass before the flow can be considered to lie in the time-periodic regime; in this case the streamlines are captured at times when the eddy shape is closest to the steady state case $h_1 = 0$, this being the reason for the slightly different starting times for the three computations shown. The time-dependent variation of channel thickness leads to periodically increasing and decreasing eddy formation in the valley of the corrugated lower surface. As the solution proceeds, material from within the initially closed eddy region is exchanged with the bulk flow, as the separatrix detaches from the right-hand side of the valley and fluid is entrained into the valley; at the same time fluid is ejected from the eddy into the bulk flow. In all three cases material exchange takes place in both directions, from the bulk flow to the eddy region and vice versa. For the case $h_1 = 0.01$ the effect appears marginal, while

for increasing amplitude h_1 the amount of material exchange grows significantly. The above results exhibit a variant of the "turnstile lobe effect", as observed for instance by Wierschem and Aksel [276] and Horner et al. [129], and therefore represent a qualitative difference to the case of Fig. 6.32. This leads to the assumption that the amount of mass exchange between the steady-state eddy and the bulk flow depends proportionally, involving a geometry-dependent constant, on the energy induced by the perturbation signal, e.g. by the corrugation of the upper surface; while a certain minimum energy is necessary to initiate the process at all. Revealing in this context would certainly be that of a study of different upper corrugated surface wavelengths in relation to the lower surface contour. Furthermore, for the small Reynolds numbers considered in this example, i.e. $\mathrm{Re} \leq 10$, the influence of the nonlinear terms that have been effectively neglected are anticipated to increase slightly any effects due to a more pronounced asymmetry of the eddy.

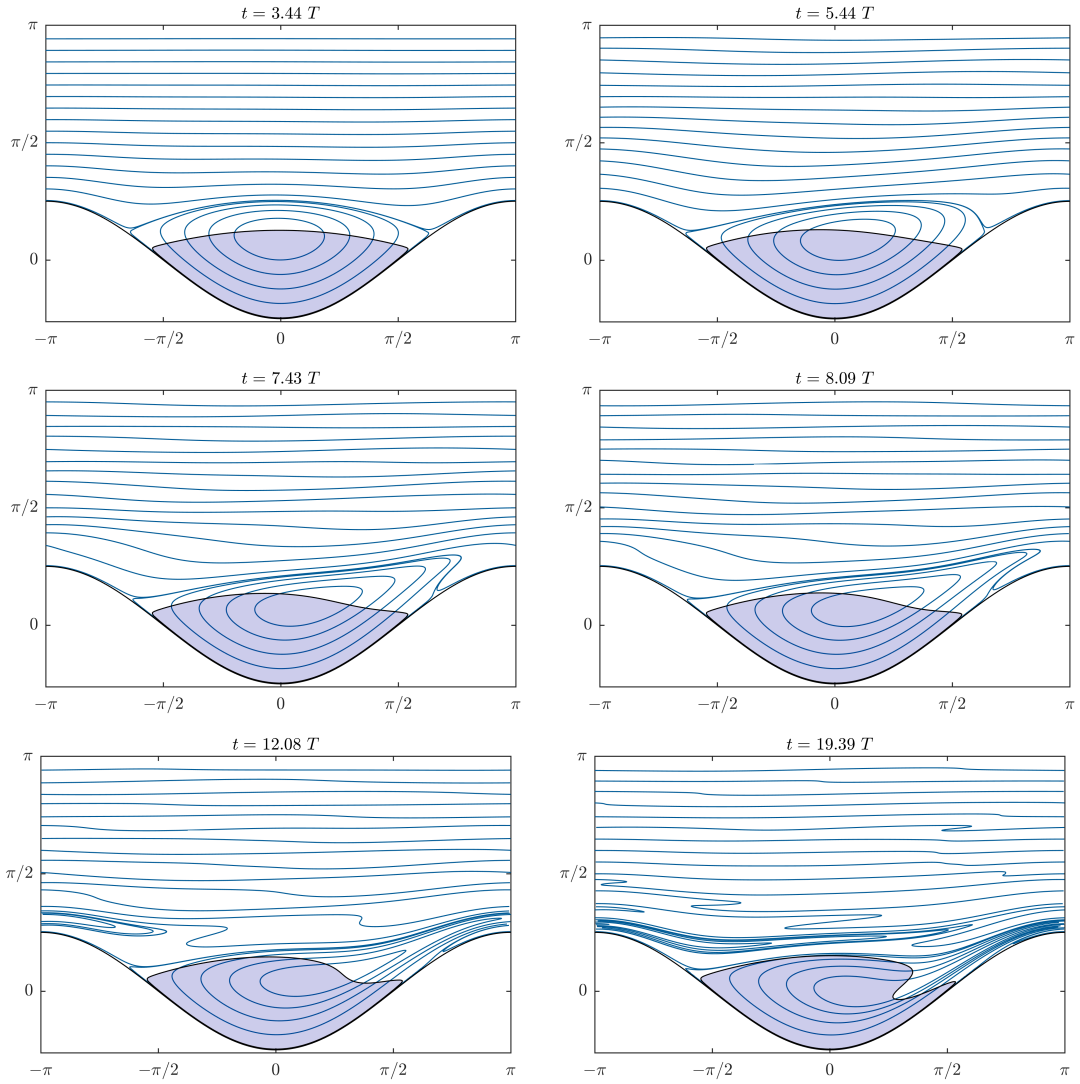


Figure 6.32: Sweep lines at different times for the case: $a = \pi/4$, $h_0 = \pi$, $h_1 = a/20$ and Re = 10; six snap-shot solutions obtained over 20 time periods. The initial shape of the sweep lines at a state with maximum eddy size ($t = 3.44\,T$) correspond to the instantaneous streamline pattern; the shaded area represents the material time evolution of the area defined at $t = 3.44\,T$ from the eddy shape of the corresponding steady-state flow.

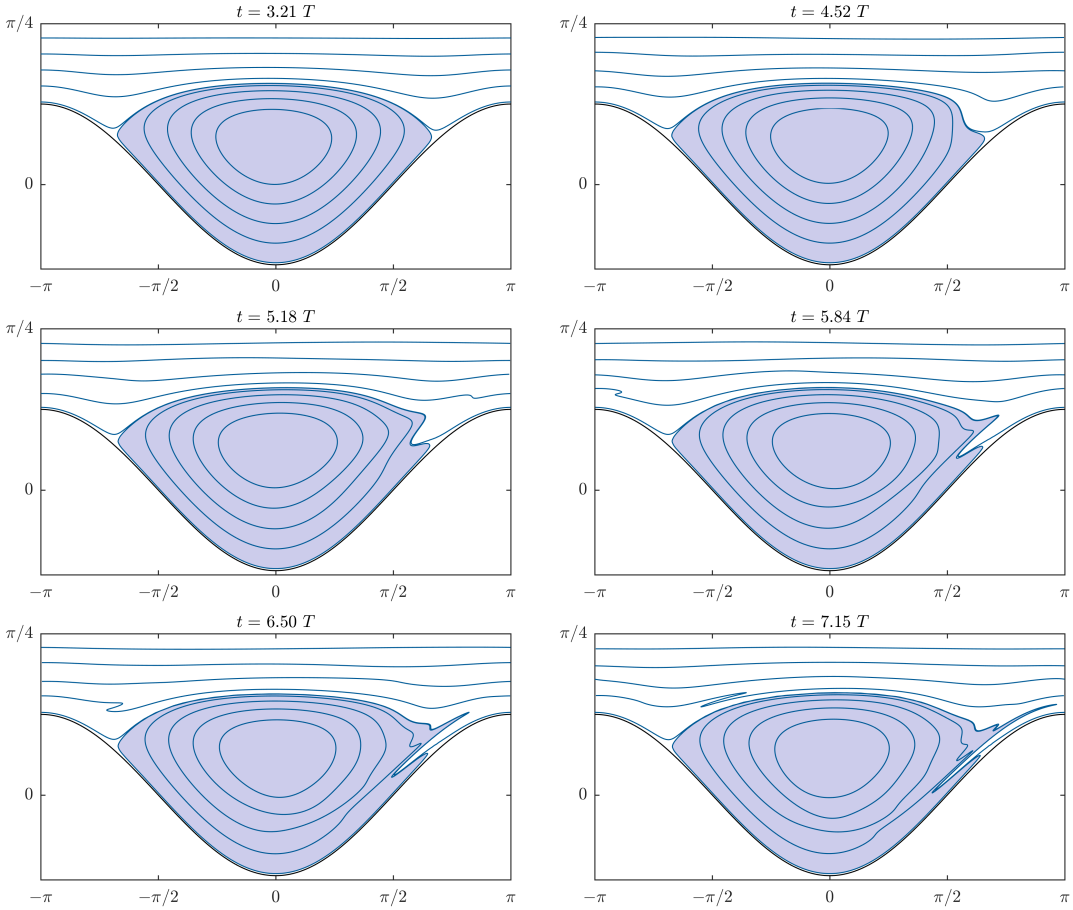


Figure 6.33: Sweep lines at different times for the case: $h_0 = \pi/4$, $a = \pi/8$, $h_1 = 0.01$, $\mathrm{Re} = 1$; the shaded area represents the particle distribution at successive times, that was initially confined within the steady-state closed eddy (upper, top left-hand image) at the onset of the unsteady behaviour. Mass transfer takes place between the eddy region and the bulk flow and vice versa.

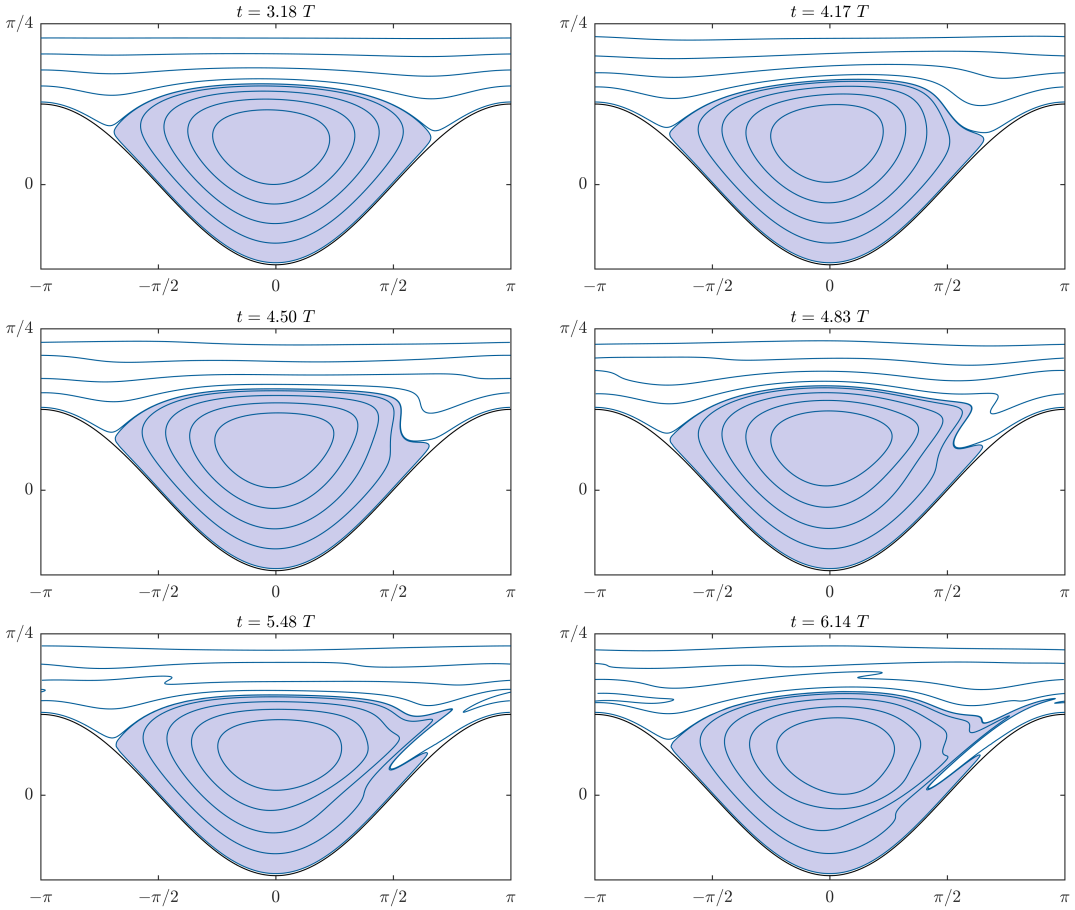


Figure 6.34: Sweep lines at different times for the case: $h_0 = \pi/4$, $a = \pi/8$, $h_1 = 0.03$, $\mathrm{Re} = 1$; the shaded area represents the particle distribution at the times shown that were initially contained within the steady-state closed eddy at the onset of unsteady behaviour - see, uppermost top left hand flow pattern. Mass transfer takes place between the eddy region and the bulk flow and vice versa.

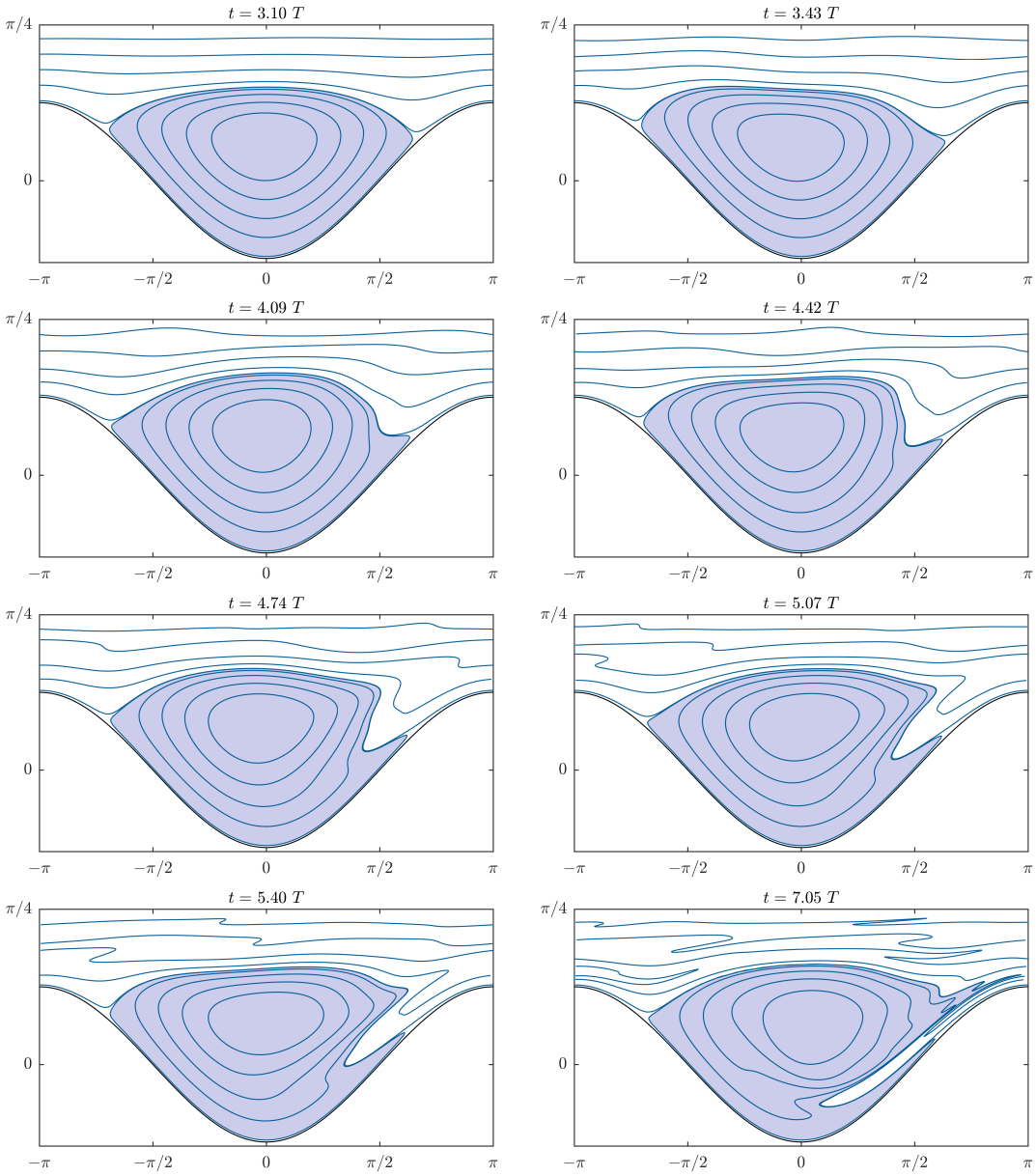


Figure 6.35: Sweep lines at different times for the case: $h_0 = \pi/4$, $a = \pi/8$, $h_1 = 0.09$, $\mathrm{Re} = 1$; the shaded area represents the particle distribution at the times shown that were initially contained within the steady-state closed eddy at the onset of unsteady behaviour - see, uppermost top left hand flow pattern. Mass transfer takes place between the eddy region and the bulk flow and vice versa.

7 Conclusions and future work

The work contained within this thesis contributes significantly to the topic and use of potential field approaches in fluid mechanics, in particular viscous incompressible flow. This is considered from two standpoints, with use being made of the classical Clebsch and Goursat approaches to derive both a Lagrangian formalism for, Chapter 2, and a first integral of, Chapter 3, the NS equations. To fully appreciate the significance of the outcomes of the thesis one needs to view and understand them in the context of the historical development of both approaches.

The Clebsch approach is rooted in the treatment of 3D inviscid flow and can be viewed as a formal extension of potential theory, where the velocity field is assumed to be given by $\vec{u} = \nabla\phi$, with Bernoulli equation arising as a natural consequence and which was indeed the goal of Clebsch's original work [64]; at the same time his velocity representation enabled a useful variational derivation of the flow equations thus contributing to the historical debate of how to embed fluid mechanics in the Lagrange formalism. The subsequent desire to extend this to viscous flow is obvious and represents the first two major achievements of the thesis:

- A new avenue to introducing a description of viscous fluid flow in the Lagrange formalism based on a novel extended discontinuous Lagrangian is established [226, 232]; deviations from classical theory are shown to admit an interpretation of the new equations of motion beyond thermodynamic equilibrium. Although not explored further, such a Lagrangian is useful for the creation of variational solution methods which normally exhibit advantageous features, from a numerical point of view and the associated analysis of its convergence properties etc., compared to pure Galerkin methods.

- A generalised Clebsch transformation for viscous flow is formulated [225], enabling the construction of a first integral of the NS equations written in terms of a generalised Bernoulli equation complemented by convection-diffusion-like equations for the potentials involved. These equations are solved for a well-known stagnation-point flow but they could be readily solved within an appropriate numerical framework and applied more widely.

In contrast, the Goursat approach has its roots in the treatment of steady Stokes flow, based on its biharmonic streamfunction formulation, and while restricted to 2D flow provides a general solution (a second integral) in terms of analytic functions, thus allowing for a reduction of the problem dimension. An analogous complex-valued formulation for 2D inertial flow does not allow for the construction of an explicit general solution but was shown by [199] to facilitate (much more complicated) implicit parametrized solutions. As above, the desire to extend these ideas to unsteady 3D viscous flow with applications is obvious, leading to the remaining two major achievements of the thesis:

- The inclusion of unsteadiness and/or moving to three dimensions required a fundamental abstraction of the above ideas:
  - firstly, an existing first integral for 2D viscous flow [230] was identified as a generalisation of the classical Goursat approach enabling it to be written in a compact form [171];
  - secondly, the introduction of an additional potential field enabled the incorporation of unsteadiness [172];
  - thirdly, an exact first integral of the unsteady 3D NS equations, based on the introduction of a real-valued tensor potential field, was derived [231]. The resulting gauge freedoms allow for a versatile tuning of the equations for different fields of application. A corresponding first integral of the dynamic boundary condition, present in free surface flow applications, has been established. The usefulness of the methodology is established via the solution, analytically and numerically, of a hierarchy of well-known benchmark flow problems.

- A new FEM for the solution of 2D and axi-symmetric 3D flows, based on the first integral of the NS equations and particularly suited to free surface problems, has been constructed which incorporates an accompanying suitably customised and efficient algebraic multigrid solution strategy. The method is theoretically analysed and numerically validated prior to solving a number of challenging thin film flow problems of topical engineering interest.

Note that, much of Part II remains to be published, the exception being the work reported in Sec. 6.4 [172]. A directly related supplementary article that has appeared is [232].

The above key outcomes are reviewed in greater detail below, following which an outlook as to how they might be extensively utilised and extended further is provided.

7.1 Summary and discussion of results

A direct comparison between the relative merits of the complementary Clebsch and Goursat approaches followed in Part I of the thesis is not feasible at the present time, plus it remains an open question as to whether the two can ever be merged in a unified framework. The former aims at a description of viscous flow based on a Lagrange formalism by thermodynamic means, leading at present to purely theoretical results and leaving many open questions with respect to the related physics. By comparison the latter, a purely mathematical integration procedure centred on tensor calculus, offers much greater scope, utility and applications wise, such as in the framework of the numerical method developed and employed in Part II to solve a range of thin film flow problems of engineering interest.

Clebsch-related potential approaches

Section 1.2.2 highlights that in contrast to Lagrangian variables the more common fluid description in terms of Eulerian variables poses severe problems for the establishment of a variational principle for viscous flow. Early investigations by Millikan [178] and later Finlayson [99] proved that such a variational principle purely in terms of primitive variables is impossible and thus a representation in terms of auxiliary potential fields required. Furthermore, Scholle [218] provides a general explanation for the necessary use of potential representations of the observables for different physical systems along the lines of a rigorous analysis of the fundamental symmetries the Lagrangian has to fulfil, with particular regard to Galilean invariance.

Eulerian principles have been proposed for the particular cases of Stokes flow by Helmholtz [124] and for ideal fluids by Clebsch [64], Lin [162] and Seliger and Witham [236], the latter based on a Clebsch-like velocity representation, equation (1.70). By comparing the latter with the original Clebsch formula (1.23) for isothermal flow, shows that any kind of extension of the system, by additional degrees of freedom as well as by additional physical effects, requires adjustment of the potential representation, see e.g. Wagner [271]. The argument of the physical insignificance of such a potential based variational principle is opposed to Salmon's [213] remark that all the potentials are more or less related to either thermodynamics or at least to "particle labels" of the Lagrangian fluid description.

Based on the above preliminary findings, Scholle [222, 223] suggested the Lagrangian given by equation (2.4) for viscous flow by suitably supplementing the Lagrangian (1.68) of [236] with additional terms from the Helmholtz principle (1.51), fulfilling all the requirements given in [218]. In Chapter 2, Sec. 2.1, this Lagrangian is analysed revealing it to be only partially successful: while the phenomenon "viscosity" occurs in a qualitatively correct manner the equations of motion resulting from the variation of Hamilton's prin-

ciple differ significantly from the NS equations, including their differential order and the occurrence of a further physical degree of freedom, the thermasy. Similar differences have also been found in the case of compressible flow with pure volume viscosity by Zuckerwar and Ash [288], and the appearance of an additional physically relevant degree of freedom, in particular, appears also in the variational formulation of heat conduction proposed by Anthony [11].

An unexpected feature of the modified equations is that the momentum density resulting as a canonical Noether observable, does not equal the mass flux density. The difference between both, termed *quasi-momentum density*, needing to be explained physically; according to [218], the quasi-momentum density could be due to contributions to the system's momentum balance beyond the scope of the continuum hypothesis on a molecular scale, e.g. Brownian motion. This hypothesis is tested by means of several "benchmark problems" which, in some cases, provide qualitative agreement with the classical NS equations including viscosity effects, but in other cases reveal the occurrence of strong deviations and partly non-physical behaviour.

As a remedy for the observed problems and based on an analogy between quantum mechanics and fluid mechanics the Lagrangian (2.4) has been refined, leading to a new *discontinuous* Lagrangian (2.32). In a first step, calculation of the corresponding Euler-Lagrange equations required the abstract extension of Lagrange formalism to discontinuities which is investigated in detail in Sec. 2.2. Next, Secs. 2.3-2.4, it is proven via a careful analysis that the dynamics resulting from Hamilton's principle based on equation (2.32) can consistently be interpreted as a generalisation of the theory of viscous flow towards thermodynamic non-equilibrium, with recovery of the classical NS-equations and the balance of inner energy when applying the limit $\omega_0 \to \infty$ to the resulting equations of motion. The physical dimension of ω_0 is a reciprocal of time and can be interpreted as a relaxation parameter which is related to the frequency of a "stick-slip" friction mechanism between fluid layers. As a striking feature, the application of the limit $\omega_0 \to \infty$ *directly* to the Lagrangian (2.32) fails; indicating that a variational formulation of viscous flow cannot be achieved using a continuous Lagrangian.

It is important to note that the Euler-Lagrange equations resulting from the two Lagrangians (2.4, 2.32), the latter one derived in Sec. A.1.3 of Appendix A.1, can also be interpreted as a *first integral* of the equations of motion; thus providing a generalisation of the methodology of Sec. 1.2.1(a), indicating that the use of potential fields seems inevitable for the construction of first integrals of the equations of motion in fluid mechanics.

Closely related to the above new Lagrangian, which involves the Clebsch variables from the ideal flow principle [236], is the question of whether the Clebsch transformation itself can be extended beyond its classical inviscid flow context. This is investigated in Sec. 2.5.

The essential problem inhibiting the application of the Clebsch transformation to viscous flow is due to the friction force density $-\nu\Delta\vec{u}$ in the NS equations which does not seem to fit into the scheme (1.28). The problem of finding a decomposition of the form (1.28) with prescribed Clebsch variables α, β is handled for an *arbitrary* vector field $\vec{a}$ using a new approach. Consequently, a generalised Clebsch transformation is developed which also covers the case of incompressible viscous flow.

Although the idea presented here follows similar objectives to [171, 230], the methodical approach involving a generalized Clebsch transformation differs significantly and results in a formulation which requires fewer potential fields in the 3D case; a notable benefit. Depending on the individual flow problem, the equations may be simplified further by utilising the combined gauge transformation (1.24) and (2.63) with functions $f(\alpha, \beta, t)$, $g(\alpha, \beta, t)$ and $h(\alpha, \beta, t)$ fulfilling (1.25, 1.26) and $F(\alpha, \beta, t) = -\partial f/\partial t - g\partial h/\partial t$. Associated boundary conditions for the Clebsch variables can be derived straightforwardly from the known boundary conditions for the observable fields, e.g. from the no-slip/no-penetration condition at a solid surface at rest as: $\vec{0} = \vec{u} = \nabla\phi + \alpha\nabla\beta$.

Goursat-related potential approaches

The principal aim of Ch. 3 was to derive a first integral representation of the full unsteady incompressible NS equations, for use as an alternative starting point for the solution of 3D viscous flow problems; in contrast to previous work the construction of the first integral does not rely on Clebsch-like velocity representations but on a generalisation of the Goursat-related approach as demonstrated in Sec. 1.2.1(b). Although representing a novel achievement in itself, not unexpectedly the emphasis was broadened to encompass a number of related topics; these have been explored and reported in tandem, in some cases representing a future research area in its own right and thus left as such with a constructive way forward having been provided.

The investigation starts with a generalisation of the methodology of Sec. 1.2.1(b) to 2D unsteady flow using complex variables, Sec. 3.2; the corresponding representative formulas are particularly well suited for spectral (i.e. Fourier) solution methods as used in Ch. 6. A central finding is that the incorporation of unsteady effects is necessarily accompanied by a more complex potential representation: instead of two real-valued potentials (Ψ and Φ), in the unsteady case one complex-valued and one real-valued field (χ and Ψ) are necessary. In tandem with the field equations (3.12)-(3.13), a first integral of the dynamic boundary conditions present in free surface applications is established; which, in the steady case, reduces to a simple pair of Dirichlet and Neumann conditions (3.28), (3.26) for Φ, representing a key feature for the development of an efficient free surface FE solver as described in Ch. 5.

In Secs. 3.3 - 3.4 the first integral in its full 3D and unsteady form is derived in an analogous fashion to Maxwell's use of potential fields in developing his classical electromagnetic theory and governing equations [133]. A tensor potential is introduced as an auxiliary unknown allowing the NS-equations to be recast as the divergence of a tensor quantity set to zero. Integration leads to a tensor equation that splits conveniently into symmetric and skew-symmetric parts. Following this it is shown that the gauge freedoms present can be exploited in an astute way leading to a re-ordering of the first integral via the elimination of mixed derivatives; this results in a more tractable equation set consisting of a vector-valued linear inhomogeneous diffusion equation and a tensor-valued generalised Poisson equation possessing the distinguishing feature of reduced non-linearity for both unsteady and steady flows. Furthermore, traceless forms of the same are derived, leading for unsteady (steady) flow to just eight (six) independent PDEs for eight (six) unknowns. Steady Stokes flow leads to a further reduction still, to simply four independent PDEs for four unknowns.

The inviscid (zero viscosity) limit of the first integral is investigated showing that, starting with its re-ordered form, the Euler equations are recovered proving that it satisfies this important subset together with the requirement that energy, momentum (in the absence of external forces), angular momentum and helicity are all conserved. In addition, for the case of steady flow it is shown in the context of finding a variational formulation how the first integral can be used to define a Lagrangian enabling it to be written in a self-adjoint form which can be useful in relation to representing particular flow problems.

Starting with the first integral, in Sec. 3.5 three well known 3D classical benchmark viscous flow problems are solved. The boundary conditions required to do so are defined and particularly those to be applied at a free surface are derived in full with the dynamic condition itself taking the form of a first integral. Two of the problems investigated are amenable to analysis – that of (i) a translating disc in a viscous fluid and (ii) a non-axisymmetric stagnation flow. In both cases the new approach leads to a non-conventional but straight forward solution procedure yielding results consistent with counterparts available in the open literature. In addition, for the translating disc problem the well-known potential representation of Papkovich and Neuber [185], known from linear elasticity theory, is reproduced; validating the calculations carried out and demonstrating the use of such a general formulation as a versatile means of representing viscous flow.

The third problem, that of viscous flow in a cubic domain, is considered from two perspectives: (iii) as a classical lid-driven cavity; (iv) from the point of view of evolving flow in periodic geometries synonymous with the DNS of viscous flow. Both situations require a numerical approach to solve them. Since (iii) involves the satisfactory use of a discrete version of the equation set defining the first integral – finite difference, volume or

element, any one of which will suffice – attention was directed at this problem for validation purposes. A finite difference methodology was used to obtain solutions for three different Reynolds number flows, yielding results in very good agreement with, for example, the corresponding predictions of Ding et al. [86]. A satisfactory outcome in itself, but just as importantly the auxiliary boundary conditions for the tensor potential entries, derived in Section 3.5.1(c), are shown to confirm the establishment of a system of equations that are uniquely solvable. For (iv) the traceless form of the first integral proves to be extremely beneficial since the pressure, in addition to typically causing regularity problems in the numerical treatment [260] of evolving flows in non-periodic geometries, is not involved pointing to an alternative formulation. By Fourier decomposition a set of uniquely solvable quadratic equations for the coefficient functions is obtained, describing the time evolution of the flow and therefore a promising starting point for future exploration.

Finally, in Sec. 3.5.5 a methodology for the treatment of film flows over non-planar surfaces in the presence of symmetries is provided which is used later on in Sec. 6.3 for the problem of thin film coating of a hemisphere. For this purpose the previously introduced variational principle from Sec. 3.4.3 proves useful because the Lagrangian (3.74) is easily amenable to coordinate transformation. As an outcome, the advantages found for the 2D Cartesian case, i.e. that the equations of motion can be described in terms of two real-valued auxiliary potentials which allow for a convenient form of the dynamic boundary condition, extend to 3D domains with an arbitrary symmetry; for example symmetric flows described in cylindrical or spherical coordinates. The Cartesian 2D case is recovered as a special case.

A new FE flow solver based on the first integral

In Chapter 4, a theoretical description of a new FE method based on the first integral for the case of the 2D steady NS equations is given. The method is centred on a least-squares FE framework which essentially requires the set of equations to be of first order to allow for practical function spaces; the corresponding reformulation is presented in Sec. 4.2. As a basis for further convergence analysis, which is mostly confined to the linear Stokes case, a priori estimates for the analytic solution of the problem are necessary, that is estimates of the solution against the right-hand side of the field equations as well as the boundary conditions are sought. For a rather general class of elliptic boundary value problems, the present one included, such estimates can be obtained from the well-known theory of Agmon, Douglis and Nirenberg (ADN) [2], [3], which is summarised in Appendix B.2 and applied to the first integral formulation in Sec. 4.3.1. The specific norms involved in the ADN-estimate restrict the choice of norms for the least-squares minimisation: in the present case the first integral system is classified as an elliptic system of Petrovskii

type which proves to be particularly useful as it allows for an efficient minimisation in the $L^2(\Omega)$-norm.

A weak formulation based on the minimisation of a least-squares energy functional is presented in Sec. 4.3.2; here, the ellipticity of the corresponding bilinear form is demonstrated also and the existence of a unique weak solution derived via the theorem of Lax-Milgram. In this context the question arises whether boundary conditions are either weakly incorporated in the least-squares functional or strongly incorporated via a restriction of the finite element space. Although the first approach allows for a simpler analysis in the sense that it results in a conforming FE discretisation, the second approach is favourable as one primarily deals with Dirichlet type boundary conditions which are most easily treated in an explicit way; as a consequence, neither weighting nor the construction of the otherwise occurring boundary integrals is necessary. However, this comes at a price, namely the precise arrangement of the domain discretisation, especially near the boundary, enters the convergence analysis. At this point it becomes clear that a higher order FE domain discretisation should be accompanied by a corresponding high order boundary approximation in order to retain the expected convergence rate; this is, for instance, ensured by iso- or superparametric approaches. The subsequent verification of optimal $H^1(\Omega)$-convergence in Sec. 4.3.3(a) mainly relies on the ADN-estimate and an approximation theorem which estimates the principle error of interpolation in the associated FE space. Optimal $L^2(\Omega)$-convergence, see Sec. 4.3.3(b), is technically more involved and utilises convergence and regularity results from the Galerkin discretisation of a system of second order PDEs which is naturally derived from the first order one (App. B.3). Possible problems and an extension of the convergence analysis to non-smooth domains with corners are discussed in Sec. 4.3.4 and finally the full non-linear problem is considered in Sec. 4.4. From the analysis as a whole it can be concluded that optimal convergence rates, which are already known for a number of different FE discretisations for different formulations of the NS equations [107], are equally well obtained for the present least-squares first integral setting.

Subsequent to the theoretical results obtained above, Chapter 5 focus on the numerical aspects of an efficient implementation providing details of the linearisation process, least-squares weighting, matrix structure, conditioning and solution of the linear systems. Initially, implementation details of the full non-linear problem are given in Sec. 5.1 providing the resulting matrix structure, split into linear contributions and those resulting from Newton-linearisation depending on least-squares weighting factors. Firstly, the case of Dirichlet velocity and potential boundary conditions is considered, then separately the case of periodic boundary conditions which is relevant for the investigation of the shear and film flow problems investigated in Chapter 6. In this context a simple free surface iteration scheme is presented.

For the construction of the system matrix and right-hand side an isoparametric concept of triangular and quadrilateral elements of first and second order is applied; that is, the matrix is constructed element-wise and the individual integrals are approximated by element transformation and efficient numerical quadrature. Here, it is also argued why isoparametric elements fulfil the discretisation Assumptions 4.6, necessary to obtain the desired convergence results. In Sec. 5.2 conditioning of the linear systems is discussed and found to be of order $\mathcal{O}(h^{-2})$, similar to that of a standard Laplace matrix.

The method is thoroughly validated in Sec. 5.3 by checking the theoretically predicted convergence rates for different elements, grid types and weighting strategies for several examples; these involve possible complicating features such as multiply connected domains, irregular boundary points and various combinations of boundary conditions. Generally, the results are found to be in agreement with the convergence theory of Chapter 4; also the condition numbers are checked and found to behave as expected. Potential problems with mass conservation are addressed and resolved by a combination of proper least-squares weighting and the use of higher order elements. It was found that an overweight of the continuity equation by a factor between 100 and 1000 versus 1 for the other equations, in combination with at least second-order elements for all test and solution spaces, results in a sufficient level of accuracy in relation to mass conservation even for comparatively small numbers of elements.

As an ultimate test case, convergence of the Newton iteration and error behaviour for increasingly refined grids is checked for the well-known benchmark lid-driven cavity problem and compared to the solutions obtained using a standard formulation in terms of primitive variables. It is found that the optimal convergence rates obtained for the linear Stokes case also translate to the linearised equations in the Newton discretisation of the full non-linear problem; in this respect the first integral form and the primitive variable form of NS equations are comparable. The importance of ensuring adequate mesh refinement in the corners forming the solution domain is highlighted; the consequence otherwise is that the convergence rate suffers. Moreover, the reduced non-linearity of the first integral form motivated an investigation of the Newton convergence speed for a wide range of Reynolds numbers which was found to be comparable to the primitive variable case; the difference being that the first integral formulation exhibits a slightly wider radius of convergence.

One of the significant advantages of the new method, compared to classical ones, is the beneficial structure of the accompanying symmetric and positive definite systems, avoiding the complication of having to solve saddle point problems [28]. It is clearly shown that the solution of these systems can be achieved in an efficient and scalable way using multigrid techniques: especially the algebraic multigrid (AMG) approach allows for the treatment of linear systems resulting from large, sparse and *unstructured* grids. An adapted AMG

approach for the first integral formulation is proposed and thoroughly tested in Sec. 5.4. It is found that a suitable choice of the standard multigrid components, i.e. the coarsening, interpolation, smoothing and cycling procedures, allows for a virtually optimal (linear) scaling of the computational effort with the number of unknowns; this depends both on the Reynolds number and the choice of least-squares weighting.

Engineering applications

In Chapter 6 several problems of engineering interest are solved using the above solution methodology. These include the investigation and minimisation of friction between parallelly moving corrugated surfaces, material exchange between entrapped vortex regions and the overlying bulk flow in a similar lubrication setting, gravity-driven film flow over corrugated planar and over curvilinear substrates. These applications demonstrate the flexibility and efficiency of the proposed method, even in the presence of periodic and free surface boundary conditions.

A possible way to reduce hydrodynamic friction in a system consisting of lubricant oil confined between two rigid surfaces translating relative to each other is contouring of the normally flat surfaces with corrugations allowing for the controlled generation of eddies which act like hydrodynamic roller bearing [217]. The impact of such contouring on the friction developed is investigated for an idealised system of two-dimensional steady Couette flow, Sec. 6.1, and a friction coefficient, (6.2), is defined which measures the decrease of friction in a corrugated surface system compared to a flat surface system. The friction coefficient κ can be calculated conveniently from the auxiliary potential fields involved in the first integral FE formulation without the need to approximate velocity derivatives in a post-processing step as would be necessary based on a primitive variable formulation.

For three models of increasing complexity, i.e. a lubrication approximation, Stokes flow and solution of the full NS equations (in its first integral form), the effect on κ is investigated in terms of the geometric parameters of the contoured surface. Also Reynolds number effects are investigated and generally the validity of the two reduced models checked. It is observed that the lubrication approximation, for the gaps h_0 under consideration, underestimates the friction significantly while, for $\mathrm{Re} < 100$, the Stokes solution gives a good approximation. Given h_0 as a technical constraint, a comparison of different surface contours – resulting in the same fluid volume in the gap, which is ensured by variation of the amplitude – reveals a theoretical optimum for a castellated shape. However, if the surface contour exhibits increasingly sharp ridges – when the geometry is varied from a sinusoidal contour in the direction of a castellated contour – the Reynolds number effect on κ becomes important also and can no longer be neglected. The latter provokes a quicker increase in friction for sharper ridges compared to smooth harmonic contours, so that for

a certain critical Reynolds number (which is outside the considered range) a trade-off is to be expected.

In Sec. 6.2 the first integral FE formulation of the full steady 2D-NS equations presented in Chapter 5 is used to solve for gravity-driven film flow down an inclined corrugated substrate; in doing so the specific remarks of Sec. 5.1.2 for periodic and free boundaries are implemented. The FE method is complemented by a spectral Fourier method based on the complex-valued representation of the first integral from Sec. 3.2. The latter is, in contrast to the generic FE method, tailored to periodic film flow applications and more suitable for possible stability analyses.

In order to verify both methods, the results obtained are compared to well-known experiments by [234, 278] and found to be in excellent agreement. Further studies investigate the impact of varying inclination angle, Reynolds number and Capillary number on the free surface shape and the internal flow structure, when the substrate contour is either sinusoidally or box-shaped. While the LSFEM delivers accurate results and streamlines for a wide range of parameters, correctly reflecting the complex interplay between the two non-linearities due to convection and curvature, the series solution of the spectral method is observed to diverge for streamlines approaching the contoured bottom surface; leading to inaccurate resolution of eddies in the vicinity of the corrugations and generally to streamlines which are only accurate close to the free surface. Similar behaviour was observed by Scholle [219] but seems unproblematic as long as only the free surface shape is sought. However, it is found that for free surface shapes which deviate from a harmonic shape too much, the error of the spectral method is still comparatively large; this could be due to the Taylor approximation (6.36) or due to conditioning problems being typical for such Fourier approaches.

In addition to film flow over planar corrugated surfaces, an outlook as to the treatment of coating flow over curvilinear surfaces is also given in Sec. 6.3, based on the variational description induced by the Lagrangian (3.74). In particular, coating flow over a hemisphere is investigated and, although the methodology of Chapters 4 and 5 could equally well be extended to spherical coordinates (compare formulas (3.136)-(3.139) with (3.142)-(3.145)), considerations are confined to the lubrication approximation and an asymptotic analytic solution for very thin films. The corresponding result for the film thickness (6.65) corresponds to that of Takagi and Huppert [246], as demonstrated in Appendix D.2.2 while, differently, the velocity profile (6.61) is not exactly parabolic as assumed by [246] nor is it self-similar.

As an important outcome of the latter study, it is found that all the advantages present for Cartesian 2D flow, i.e. the description of the first integral equations in terms of only two scalar quantities (Ψ and Φ) and the dynamic boundary condition as a pure gradient

condition on Φ, similarly extend to 3D axisymmetric film flows; this broadens the range of applicability significantly.

A particular aspect of lubrication-like flows involving corrugated surfaces is the process of material exchange between the fluid inside the valleys of the corrugations and the bulk flow which, in some applications, is desirable to be controlled for instance to achieve a homogeneous mixture of the lubricant. This is investigated in Sec. 6.4 using an unsteady 2D Couette flow model in a channel bounded by two sinusoidally corrugated surfaces. Calculations are carried out by means of a semi-analytical and a purely numerical method which are adapted, in comparison to Chapter 5, to deal with the time-dependency.

For the analytical approach a generalised Goursat form with a third holomorphic function is established for the streamfunction, which more generally can be written in the form: $\psi = \mathrm{Re}\left[g_0(\xi) + \bar{\xi} g_1(\xi) + \mathrm{Re}\,\bar{\xi}^2 g_2(\xi)\right]$. It is noted that although this elegant extension of the classical complex variable method is not generally applicable, it can at least be applied to flow problems with negligible non-linear inertial effects and small Reynolds numbers taking into account the remaining unsteady effects. Asymptotic analysis reveals a time-dependent eddy structure invoking a material exchange, a topical subject of research in relation to coating and film flows. Via numerical studies, based on implicit Crank-Nicolson time discretization in combination with a weak Galerkin FE method, this material exchange is revealed in detail by visualisation of sweep lines, showing the associated flow structure via a sequence of snap-shots of the ensuing motion. As a main outcome of the investigation it is concluded that the observed "turnstile lobe" effect is induced purely geometrically, even when inertial effects are absent, and can be controlled by the relative speed of the surfaces and the geometrical properties of the corrugation contour.

7.2 Suggestions for future work

The above summary and discussion of results demonstrate that the research begun in this thesis can be related to a number of very different fields, such as variational calculus and particularly the classical Lagrange formalism with possible extensions to discontinuities involving also distribution theory. Certainly the stick-slip friction law inherent in the new variational formulation for viscous flow is somehow equivalent to a stochastic representation (allowing for an analysis from a completely different viewpoint). Related to variational principles is the Noether theorem, symmetry analysis and gauge theory in the broadest sense. All the research involving Clebsch-like representations is also deeply related to topological fluid mechanics and some of the findings would most likely have their natural description in the field of differential geometry. The more applied FE related

analysis is based on PDE theory, a wide spectrum of functional analysis and not least numerical methods.

Considering this, the thesis provides a number of important novel ideas, sketches possible utilisation of the same and in particular tries to cross-link the ideas with a focus on overview. Although the research presented is clearly formative many avenues to further development are raised throughout the chapters; a selection of specific avenues for future research are mentioned explicitly below.

Certainly, the discontinuous variational principle for viscous flow presented in Chapter 2 requires further verification. The question arises as to whether the occurrence of discontinuous interfaces inside fluid flow is an artefact of the model or if such phenomena really exist on a microscopic scale. Although for large ω_0 the discontinuities are physically reduced to fluctuations on a micro scale, it would be of great interest future work wise to explore further the dynamics of the discontinuous interfaces and the induced physical effects beyond thermodynamic equilibrium for various flow geometries. It is further noted that the approach is deeply related to the stochastic variational formulation of Arnaudon and Cruzeiro [13, 14], see also [148], which is worth a detailed comparison. Here, the question arises whether a stochastical description provides tools of analysis which are otherwise unamenable. From a less methodical and more physical point of view, a comparison to Constantin et al. [71] (formula 2.12) – see also the work of Feireisl and Vasseur [94] related to a fluid mechanics model proposed by Brenner [47] – would be illuminating. In the latter a new approach to continuum fluid mechanics is proposed, based on the concept of *two* different velocities: the mass-based (Eulerian) velocity u_m derived from the classical notion of mass transport, and the fluid-based (Lagrangian) volume velocity u associated to the motion of individual particles (molecules); the resulting equations exhibit similar additional terms compared to the NS equations as observed in (2.43), (2.44) and are explained by thermodynamics.

Generally, by evaluating the dynamics induced by the Lagrangian (2.32), it has been demonstrated how Lagrange formalism applies to physical problems with discontinuities. Independent of the particular problem of viscous flow, the general formalism specified in Sec. 2.2 may also be a valuable mathematical tool for embedding various discontinuous phenomena into the Lagrange formalism, such as phase boundaries between immiscible fluids, propagating shock waves in gaseous media, flame fronts, detonation shocks and also interfaces in solids like micro cracks [79] and grain boundaries. Discontinuities also occur in some optimum control problems [120]. The extended formalism can be utilised for FE simulations of such phenomena without the imperative of considering the related matching conditions explicitly, since they result automatically from the respective Lagrangian. It is therefore realistic to expect an improvement of, for example, numerical algorithms.

First integrals of the type described in Sec. 2.5, related to the generalised Clebsch transformation, may be relevant from the point of view of initiating new simulation methods in fluid dynamics: apart from the analytic example provided in Sec. 2.5.2 the Clebsch transformation has already been proven to be useful for solving flow problems by Prakash et al. [195] in the case of inviscid flow, see also Cotter and Holm [74] and related work. There is therefore a realistic perspective for extending the method to many other viscous flow problems. It should be stressed that the generalised Clebsch transformation developed here applies to an arbitrary vector field $\vec{a}$ and can therefore be generalised to, for example, non-Newtonian fluid flows and non-conservative external forces, as well as to the Navier-Stokes-Duhem equations in order to include phenomena occurring in compressible flow such as trans- and hypersonic flows, sound and shock waves. Finally, Clebsch variables apply to dislocations in crystals [224] based on an analogy between vortices and dislocations [205], giving rise to a dynamic theory of dislocations. Based on these considerations, perspective exists for establishing a field theory of plasticity.

With respect to Chapter 3 it is clear that there remains considerable scope for further advancement, since the principle focus of the present work has been the new approach and theory underpinning the establishment of the first integral and its subsequent validation via the solution of a number of benchmark test problems; the investigation of a 3D problem involving the presence of a free-surface, such as that of thin film flow over surface topography [7, 267], represents an obvious avenue to explore. For this purpose a FE method of the form described in Chapters 4 and 5 could be established but, considering the difficulties with the numerical method in Sec. 3.5.3, the most suitable gauge for the underlying first integral equations is not obvious. However, note also that the numerical schemes suggested in Chapter 3 and particularly the LSFEM described in Chapters 4 and 5 describe only *one* possible way of utilising the new formalism while completely different opportunities exist, for instance, along the lines of Ranger [198, 199] whose work is rather unrecognised.

The new approach promises to be other than just useful for deriving different existing potential formulations from within a unified framework but able to serve as a source for further representative formulas with reference to specific applications, as sketched out for the case of unsteady flow involving periodic boundary conditions. Such formulas can serve as the starting point for both new analytical solutions and numerical techniques; in this sense the variational principle established in Section 3.4.3 points a promising way forward for further research. An interesting feature of this gauge variant is, that general 3D coating flows in the presence of symmetry can be handled in a similarly elegant way as the 2D Cartesian case. In this respect, via the full solution of (3.142)-(3.145) the calculations for flow over a hemisphere in Sec. 6.3 could be extended to full NS flow and various other

surfaces with symmetry.

Generally, the numerical method described in Chapters 4 and 5 could be optimised in various ways. First of all, a further improvement in performance is desirable, i.e. with respect to the precise multigrid components and a more sophisticated parallelisation strategy; this should be accompanied by a more detailed benchmark test involving comparisons to a standard code based on a primitive variable formulation and also to other codes based on first-order formulations of the NS equations; the latter being more natural competitors. In contrast to previous benchmark tests, what follows should focus on free surface applications, which is a major field of application for the first integral method. The most obvious extensions of the 2D code physical-wise would be multilayer flows, which can be treated completely analogously to the one-layer case; also thermal effects can be coupled to the equations in a way which is no more complicated than in classical implementations based on primitive variables. In contrast, a possible extension to non-Newtonian material laws would be more involved.

The above generalisations could also be used to refine the studies in Chapter 6 which are, so far, based on rather strong and unrealistic assumptions (steady, isothermal and incompressible Couette flow as a hydrodynamic friction model). Moreover, the remaining error in the semi-analytic film flow method of Sec. 6.2.1 should be investigated, before the method could be used as a basis for stability studies. The work on the turnstile lobe problem could be continued in several ways. In particular, future studies are planned that will include the non-linear inertial terms in the field equations and utilise the approach to explore free surface film flows making use of the elegant reformulation (3.19, 3.23) of the kinematic and dynamic boundary conditions in order to find solutions for solitary surface waves. Analysis of the stability of steady-state base flows, subject to a small disturbance away from equilibrium, is a second attractive research area for application of the field equations (3.12) and (3.13). Also, since they are in complex form they represent a suitable mathematical framework from which to calculate the time evolution of small wave-like perturbations; in combination with the complex form (3.23) of the dynamic boundary condition, the onset of surface waves in film flows [234] can be explored.

A Viscous flow and the 3D first integral

A.1 Variational principle for viscous flow - Calculations

A.1.1 Symmetries and associated Noether balances of (2.4)

In article [218] a general analysis is provided concerning the analytical structure of Lagrangians in continuum theories fulfilling invariance with respect to the full Galilei group. In the same paper a general scheme for Lagrangians is constructed. Using Noether's theorem, canonical formulae give rise to the identification of the relevant observable fields like mass density and flux density, momentum density, stress tensor, energy density and Poynting vector.

The analysis given in [218] is rigorously applied to Lagrangian (2.4): for simultaneous invariance with respect to time and space translations and Galilei boosts, a collective symmetry criterion, the *duality criterion*:

$$\ell\left(\Psi^i, \overset{\circ}{\Psi}{}^i, \nabla\Psi^i + \frac{1}{t}\vec{K}^i\left(\Psi^j\right)\right) = \ell\left(\psi^i, \dot{\psi}^i, \nabla\psi^i\right) , \tag{A.1}$$

has to be fulfilled, where:

$$\begin{aligned}
\dot{\psi} &= \frac{\partial\psi}{\partial t} , \\
\overset{\circ}{\Psi} &= \left\{\frac{\partial}{\partial t} + \nabla\zeta\cdot\nabla\right\}\Psi , \\
\zeta &= \frac{\vec{x}^{\,2}}{2t} ,
\end{aligned}$$

are the conventional time derivative, the dual time derivative and the generating field, while:

$$\psi^i = K^i\left(\Psi^j, \zeta, \nabla\zeta\right) , \tag{A.2}$$

$$\vec{K}^i\left(\Psi^j\right) = \lim_{\zeta,\nabla\zeta\to 0}\frac{\partial K^i}{\partial(\nabla\zeta)} , \tag{A.3}$$

is the dual transformation and the corresponding infinitesimal generator. In the case of Lagrangian (2.4) the dual transformation takes the form:

$$\begin{aligned} \vec{u} &= \vec{U} + \nabla\zeta\,, & \varphi &= \Phi + \zeta\,, \\ \alpha &= A\,, & \beta &= B\,, \\ \varrho &= P\,, & s &= S\,, \\ \vartheta &= \Theta\,, \end{aligned} \tag{A.4}$$

fulfilling criterion (A.1) as required. One consequence of which is that the mass balance:

$$\frac{\partial \varrho}{\partial t} + \nabla\cdot(\varrho\vec{u}) = 0\,, \tag{A.5}$$

is automatically fulfilled and $\varrho\vec{u}$ is identified as the mass flux density. In the following the computation of the canonical stress tensor, energy density and Poynting vector from Noether's theorem is skipped, with the focus on the canonical momentum density which [218] takes the form :

$$\vec{p} = -\frac{\partial \ell}{\partial \dot{\psi}^i}\nabla\psi^i = \varrho\left[\nabla\varphi + \alpha\nabla\beta - s\nabla\vartheta\right]\,, \tag{A.6}$$

and need not to be identical to the mass flux density $\varrho\vec{u}$. Since the dual transformation formula (A.4) contains a real $\nabla\zeta$-dependence, the relation:

$$\vec{p} = \varrho\vec{u} + \vec{p}^{\,*}\,, \tag{A.7}$$

is given with *quasi-momentum density*:

$$\vec{p}^{\,*} = -\frac{\partial}{\partial t}\left[\frac{\partial \ell}{\partial \dot{\psi}^i}\vec{K}^i\right] - \nabla\cdot\left[\frac{\partial \ell}{\partial \nabla\psi^i}\vec{K}^i\right] = -\nabla\cdot\left(\frac{\vartheta}{T}\left[2\eta\underline{D} + \eta'\left(\nabla\cdot\vec{u}\right)\underline{1}\right]\right)\,. \tag{A.8}$$

Hence, a striking feature of Lagrangian (2.4) is the difference between mass flux density and momentum density which is in contrast to classical continuum mechanics.

A.1.2 Steady solution of (2.8-2.10) for Poiseuille flow

Considering the flow geometry $\vec{u} = u(y)\vec{e}_x$ and assuming time-independent thermasy ϑ, the evolution equation (2.10) takes the form:

$$u(y)\frac{\partial}{\partial x}\left(\frac{\vartheta}{T}\right) = 1\,, \tag{A.9}$$

which obviously implies the general solution (2.14). Making use of (A.9), equations (2.8) read:

$$\frac{\nabla p}{\nu\varrho_0} = \left\{u\frac{\partial}{\partial x} + u'\vec{e}_y\otimes\vec{e}_x\right\}\left[u'\left(\vec{e}_x\otimes\vec{e}_y + \vec{e}_y\otimes\vec{e}_x\right)\nabla\left(\frac{\vartheta}{T}\right) + \frac{\vartheta}{T}\Delta\vec{u}\right] - \frac{1}{2}u'^2\nabla\left(\frac{\vartheta}{T}\right)\,,$$

$$= uu' \frac{\partial}{\partial x} \left\{ \vec{e}_x \frac{\partial}{\partial y} + \vec{e}_y \frac{\partial}{\partial x} \right\} \frac{\vartheta}{T} + u'' \vec{e}_x + \left[u'^2 \frac{\partial}{\partial y} \left(\frac{\vartheta}{T} \right) + u' u'' \frac{\vartheta}{T} \right] \vec{e}_y - \frac{1}{2} u'^2 \nabla \left(\frac{\vartheta}{T} \right) ,$$
$$= \left[u'' - \frac{3u'^2}{2u} \right] \vec{e}_x + \frac{\partial}{\partial y} \left[\frac{u'^2}{2} \frac{\vartheta}{T} \right] \vec{e}_y .$$

Taking the component in the y-direction, it follows that:

$$p = \eta \frac{\vartheta}{T} \frac{u'^2}{2} + f_2(x) ,$$

for the pressure ($\eta = \varrho_0 \nu$) with integration function $f_2(x)$. Inserting this into the x-direction equation, leads to:

$$\frac{f_2'(x)}{\eta} = u'' - 2\frac{u'^2}{u} .$$

Since the left-hand side of the above equation depends on x only and the right-hand side on y only, both sides are constant, i.e.:

$$\frac{f_2'(x)}{\eta} = -K ,$$
$$u'' - 2\frac{u'^2}{u} = -K ,$$

with $K > 0$. Hence, the pressure can be written as:

$$p = p_0 - \eta K x + \eta \frac{\vartheta}{T} \frac{u'^2}{2} , \tag{A.10}$$

while the velocity profile has to fulfill the nonlinear ODE:

$$uu'' - 2u'^2 + Ku = 0 . \tag{A.11}$$

Via the substitution $g(y) = 1/u(y)$, equation (A.11) simplifies to the second order ODE: $g'' - Kg^2 = 0$, which after multiplication with g' becomes integrable:

$$\frac{\mathrm{d}}{\mathrm{d}y} \left[\frac{g'^2}{2} - K \frac{g^3}{3} \right] = 0 ,$$

implying the first order ODE $g'^2 - 2Kg^3/3 = C$. The integration constant, C, is evaluated at the middle of the channel, $y = h/2$, where the velocity is a maximum, $u(h/2) = u_{\max}$, and therefore $u'(h/2) = 0$, implying: $C = -2K/(3u_{\max}^3)$. Hence, after re-substituting $u = 1/g$:

$$u' = \pm \sqrt{\frac{2K}{3} \left[1 - \frac{u^3}{u_{\max}^3} \right]} u ,$$

results as a nonlinear first order ODE for the velocity profile, with the positive sign valid for $0 \leq y < h/2$ and the negative sign for $h/2 < y \leq h$. Using separation of variables, the

above ODE can be solved, with the solution for $y \leq h/2$ given implicitly as:

$$y = \sqrt{\frac{2u}{K}}\, {}_2F_1\left(\frac{1}{6}, \frac{1}{2}; \frac{7}{6}; u^3/u_{\max}^3\right), \tag{A.12}$$

$$u_{\max} = \frac{9\Gamma(\frac{2}{3})^2\Gamma(\frac{5}{6})^2}{8\pi^3} K h^2, \tag{A.13}$$

with Gaussian hypergeometric function ${}_2F_1$.

A.1.3 Euler-Lagrange equations of Lagrangian (2.32)

The Lagrangian (2.32) is based on the following fields: the velocity $\vec{u}$, the three Clebsch variables Φ, α, β and the complex field of thermal excitation χ. The associated Euler-Lagrange equations are computed according to (2.36). First, variation with respect to the Clebsch variable Φ delivers the continuity equation:

$$\nabla \cdot \vec{u} = 0\,. \tag{A.14}$$

As a consequence the identity $\nabla\cdot(\xi\vec{u}) = \vec{u}\cdot\nabla\xi$, which is fulfilled for any field ξ, is considered subsequently. Next, variation with respect to the two remaining Clebsch variables α and β lead to the transport equations:

$$\mathrm{D}_t\beta = 0\,, \tag{A.15}$$

$$\mathrm{D}_t\alpha = 0\,, \tag{A.16}$$

after simple mathematical manipulation. By variation with respect to the velocity $\vec{u}$:

$$\varrho_0\vec{u} + \varrho_0\frac{\nu}{\omega_0}\nabla\cdot\left[\mathrm{i}\ln\sqrt{\frac{\bar{\chi}}{\chi}}2\underline{D}\right] = \varrho_0\left[\nabla\Phi + \alpha\nabla\beta + \frac{1}{\omega_0}\mathrm{Im}\left(\bar{\chi}\nabla\chi\right)\right], \tag{A.17}$$

is obtained. Finally, the Euler-Lagrange equation related to variation with respect to $\bar{\chi}$ leads to the evolution equation:

$$\mathrm{D}_t\chi + \mathrm{i}\omega_0\chi = \frac{\nu}{2\bar{\chi}}\mathrm{tr}\underline{D}^2\,, \tag{A.18}$$

for the thermal excitation; variation with respect to $\bar{\chi}$ delivers the complex conjugate of (A.18).

A.1.4 Equations of motion

The Euler-Lagrange equations (A.15)-(A.18) are a first integral of the equations of motion, i.e. the latter can be obtained from their differentiation as follows. Considering the

identities:

$$\mathrm{D}_t\left[\nabla\Phi+\alpha\nabla\beta+\mathrm{Im}\frac{\bar\chi\nabla\chi}{\omega_0}\right]=\nabla\left[\mathrm{D}_t\Phi+\alpha\mathrm{D}_t\beta+\mathrm{Im}\frac{\bar\chi\mathrm{D}_t\chi}{\omega_0}\right]-\nabla\otimes\vec u\left[\nabla\Phi+\alpha\nabla\beta+\mathrm{Im}\frac{\bar\chi\nabla\chi}{\omega_0}\right]$$
$$+\mathrm{D}_t\alpha\nabla\beta-\mathrm{D}_t\beta\nabla\alpha-\frac{2}{\omega_0}\mathrm{Im}(\mathrm{D}_t\chi\nabla\bar\chi)\,,$$
$$\nabla\cdot\left[\mathrm{i}\ln\sqrt{\frac{\bar\chi}{\chi}}2\underline{D}\right]=2\underline{D}\mathrm{Im}\frac{\nabla\chi}{\chi}+\mathrm{i}\ln\sqrt{\frac{\bar\chi}{\chi}}\left[\nabla\left(\nabla\cdot\vec u\right)+\Delta\vec u\right]\,,$$
$$\mathrm{i}\mathrm{D}_t\ln\sqrt{\frac{\bar\chi}{\chi}}=\mathrm{Im}\left(\frac{\mathrm{D}_t\chi}{\chi}\right)\,,$$

and the Euler-Lagrange equations (A.14)-(A.18), the material time derivative of (A.17) reads:

$$\mathrm{D}_t\vec u-\nu\Delta\vec u+\frac{\nu}{\omega_0}\left[\mathrm{i}\ln\sqrt{\frac{\bar\chi}{\chi}}\mathrm{D}_t\Delta\vec u+\mathrm{D}_t\left(2\underline{D}\mathrm{Im}\frac{\nabla\chi}{\chi}\right)\right]=\nabla\left[\mathrm{D}_t\Phi+\alpha\mathrm{D}_t\beta+\mathrm{Im}\frac{\bar\chi\mathrm{D}_t\chi}{\omega_0}\right]$$
$$-\nabla\otimes\vec u\left[\vec u+\frac{\nu}{\omega_0}\nabla\cdot\left(\mathrm{i}\ln\sqrt{\frac{\bar\chi}{\chi}}2\underline{D}\right)\right]-\frac{2}{\omega_0}\mathrm{Im}\left[\left(\frac{\nu}{2\bar\chi}\mathrm{tr}\underline{D}^2-\mathrm{i}\omega_0\chi\right)\nabla\bar\chi\right]$$
$$=-\frac{1}{\varrho_0}\nabla p-\nabla V-\frac{\mathrm{i}\nu}{\omega_0}\ln\sqrt{\frac{\bar\chi}{\chi}}\nabla\mathrm{tr}\underline{D}^2-\frac{\nu}{\omega_0}\nabla\otimes\vec u\left[2\underline{D}\mathrm{Im}\frac{\nabla\chi}{\chi}+\mathrm{i}\ln\sqrt{\frac{\bar\chi}{\chi}}\Delta\vec u\right]\,,\qquad(A.19)$$

where the pressure p is, according to:

$$p:=-\varrho_0\left[\mathrm{D}_t\Phi+\alpha\mathrm{D}_t\beta+\mathrm{Im}\frac{\bar\chi\mathrm{D}_t\chi}{\omega_0}-\frac{\vec u^{\,2}}{2}+\bar\chi\chi+V-\frac{\nu}{\mathrm{i}\omega_0}\ln\sqrt{\frac{\bar\chi}{\chi}}\mathrm{tr}\underline{D}^2\right]=\ell_{\mathrm{E.L.}}\,,$$

given as the Lagrangian evaluated for real processes, as is well-known from the classical literature [240]. Finally, a balance for the inner energy, $c_0T=\bar\chi\chi$ is derived from (A.18) according to:

$$c_0\mathrm{D}_tT=\mathrm{D}_t(\bar\chi\chi)=\bar\chi\mathrm{D}_t\chi+\chi\mathrm{D}_t\bar\chi=\nu\mathrm{tr}\underline{D}^2\,.\qquad(A.20)$$

A.1.5 Derivation of the production condition (2.42) using distributions

Within the theory of distributions a very elegant way is provided by which the generalised formalism for discontinuous Lagrangians can be understood in terms of conventional Lagrange formalism. Under the same assumptions made in Sec. 2.2, a continuous reference Lagrangian is defined by:

$$\ell_{\mathrm{c}}\left(\cdots,\varphi\right):=\ell\left(\cdots,\varphi\right)-\sum_{n=1}^{N_S}\left[\left[\ell\left(\cdots,\varphi_n\right)\right]\right]\mathcal{H}(\varphi-\varphi_n)\,,\qquad(A.21)$$

where $\mathcal{H}(x)$ is the Heaviside function giving 0 for $x<0$ and 1 for $x\geq 0$. Equation (A.21) defines a decomposition of the entire Lagrangian into a continuous part and a sum of

discontinuities; the derivative of it with respect to φ gives:

$$\frac{\partial \ell}{\partial \varphi} = \frac{\partial \ell_\mathrm{c}}{\partial \varphi} + \sum_{n=1}^{N_S} \left[\left[\ell\left(\cdots, \varphi_n\right)\right]\right] \delta(\varphi - \varphi_n)\,, \tag{A.22}$$

where $\delta(x)$ denotes Dirac's delta function. Using this, the Euler-Lagrange expression EL_N defined according to (2.36) with respect to $\psi_N = \varphi$ can be defined across the entire domain, leading to:

$$\begin{aligned}\mathrm{EL}_N &= \frac{\partial \ell}{\partial \varphi} - \frac{\partial}{\partial t}\left(\frac{\partial \ell}{\partial \dot{\varphi}}\right) - \nabla \cdot \left(\frac{\partial \ell}{\partial \nabla \varphi}\right) \\ &= \underbrace{\frac{\partial \ell_\mathrm{c}}{\partial \varphi} - \frac{\partial}{\partial t}\left(\frac{\partial \ell_\mathrm{c}}{\partial \dot{\varphi}}\right) - \nabla \cdot \left(\frac{\partial \ell_\mathrm{c}}{\partial \nabla \varphi}\right)}_{:=\mathrm{EL}_N^\mathrm{c}} + \sum_{n=1}^{N_S} \left[\left[\ell\left(\cdots, \varphi_n\right)\right]\right] \delta(\varphi - \varphi_n)\,.\end{aligned}$$

As a consequence of the above decomposition, the first term in equation (2.38), for $\psi_N = \varphi$, results in:

$$\int_{t_1}^{t_2} \sum_{n=0}^{N_S} \iiint_{V_n} \mathrm{EL}_N\, \delta\varphi \mathrm{d}V \mathrm{d}t = \int_{t_1}^{t_2} \iiint_{V} \mathrm{EL}_N^\mathrm{c}\, \delta\varphi \mathrm{d}V \mathrm{d}t + \int_{t_1}^{t_2} \sum_{n=0}^{N_S} \iiint_{V} \left[\left[\ell\right]\right] \delta\varphi\, \delta(\varphi - \varphi_n) \mathrm{d}V \mathrm{d}t\,. \tag{A.23}$$

In order to evaluate the integrals on the right hand with delta functions, a representation of $\mathrm{d}V$ in terms of local coordinates, $\mathrm{d}V = \mathrm{d}\xi \mathrm{d}S$, is used where $\mathrm{d}S$ is the surface element of the interface S_n and $\mathrm{d}\xi$ is related to the direction perpendicular to S_n. Via the substitution $\mathrm{d}\varphi = |\nabla\varphi| \mathrm{d}\xi$, the identity:

$$\iiint_{V} \left[\left[\ell\right]\right] \delta\varphi\, \delta(\varphi - \varphi_n) \mathrm{d}V = \iiint_{V} \frac{\left[\left[\ell\right]\right] \delta\varphi}{|\nabla\varphi|} \delta(\varphi - \varphi_n) \mathrm{d}\varphi \mathrm{d}S = \iint_{S_n} \frac{\left[\left[\ell\right]\right] \delta\varphi}{|\nabla\varphi|} \mathrm{d}S\,,$$

is obtained, giving (A.23) the more convenient form:

$$\int_{t_1}^{t_2} \sum_{n=0}^{N_S} \iiint_{V_n} \mathrm{EL}_N\, \delta\varphi \mathrm{d}V \mathrm{d}t = \int_{t_1}^{t_2} \iiint_{V} \mathrm{EL}_N^\mathrm{c}\, \delta\varphi \mathrm{d}V \mathrm{d}t + \int_{t_1}^{t_2} \sum_{n=0}^{N_S} \iint_{S_n} \frac{\left[\left[\ell\right]\right]}{|\nabla\varphi|} \delta\varphi \mathrm{d}S \mathrm{d}t\,. \tag{A.24}$$

Finally, considering the identity (2.39) and the vanishing of $\delta\varphi$ at $t = t_{1,2}$, the variation of the action integral, (2.38), turns out to be:

$$\delta I = \int_{t_1}^{t_2} \iiint_{V} \mathrm{EL}_i^\mathrm{c}\, \delta\varphi \mathrm{d}V \mathrm{d}t + \int_{t_1}^{t_2} \sum_{n=0}^{N_S} \iint_{S_n} \left\{\vec{n} \cdot \left[\left[\frac{\partial \ell}{\partial \nabla \psi_i} - \vec{v}_s \frac{\partial \ell}{\partial \dot{\psi}_i}\right]\right] + \frac{\left[\left[\ell\right]\right]}{|\nabla\varphi|}\right\} \delta\varphi \mathrm{d}S \mathrm{d}t\,, \tag{A.25}$$

which implies, next to the well-known Euler-Lagrange equation with respect to φ, the production condition (2.42).

A.2 Three-dimensional first integral

A.2.1 Recovery of the 2D form as a special case

In the following, proof is given that the equations derived in [171, 172] for 2D flow uniquely result from equation (3.55) as a special case of general 3D flow. For steady flow with $\partial_t \Psi_n = 0$ equation (3.49) is fulfilled via the identity $\varphi_n = \eta \Psi_n$, while the auxiliary vector field defined by (3.57) reads:

$$A_j = \partial_k \tilde{a}_{kj} + \varepsilon_{jlk} \partial_l \varphi_k = \partial_k \tilde{a}_{kj} + \eta \varepsilon_{jlk} \partial_l \Psi_k = \partial_k \tilde{a}_{kj} + \eta u_j \,.$$

Thus, equations (3.55) yields:

$$\varrho u_i u_j + (p + U)\delta_{ij} = -\partial_k \partial_k \tilde{a}_{ij} - \partial_l \partial_k \tilde{a}_{kl} \delta_{ij} + \partial_i \left[\partial_k \tilde{a}_{kj} + \eta u_j\right] + \partial_j \left[\partial_k \tilde{a}_{ki} + \eta u_i\right] . \quad \text{(A.26)}$$

Considering now $u_3 = 0$ and a completely vanishing x_3-dependence of all fields for a 2D flow, i.e. $\partial_3 (\cdots) = 0$, it is obvious that in case of the choice $\tilde{a}_{ij} = -\Phi \delta_{ij}$ for $i, j = 1, \cdots, 2$ for the modified tensor potential the three field equations for steady 2D flow are reproduced.

For the remaining components of the modified tensor potential, on the assumption that $\tilde{a}_{13} = \tilde{a}_{23} = 0$ and $\tilde{a}_{33} = -\zeta(x_1, x_2)$, equation (A.26) is fulfilled identically for indices $i = 1$, $j = 3$ and $i = 2$, $j = 3$; whereas when the indices are $i = j = 3$ it gives:

$$\partial_k \partial_k \zeta = p + U - \partial_k \partial_k \Phi \,, \quad \text{(A.27)}$$

which is a Poisson equation for ζ and therefore solvable. Note that ζ has no influence on the other equations and therefore has no physical effect.

A.2.2 Free surface boundary conditions

Consider the kinematic and dynamic boundary conditions at a free surface. Assuming a parametrisation of the free surface in terms of:

$$x_i = f_i(s_1, s_2, t) \,, \; i = 1, 2, 3 \,, \quad \text{(A.28)}$$

the two tangential vectors $t_i^{(1)}$, $t_i^{(2)}$ given by:

$$t_i^{(\lambda)} := \frac{\partial f_i}{\partial s_\lambda}, \quad \text{(A.29)}$$

are orthogonal and normalised. Together with the normal vector $\vec{n}$, an orthonormal basis exists locally fulfilling the relations:

$$n_j = \varepsilon_{jpq} t_p^{(1)} t_q^{(2)} , \tag{A.30}$$

$$t_i^{(1)} = -\varepsilon_{ikl} n_k t_l^{(2)} , \tag{A.31}$$

$$t_i^{(2)} = \varepsilon_{ikl} n_k t_l^{(1)} . \tag{A.32}$$

The kinematic boundary condition at a free surface is given by:

$$0 = [\dot{f}_j - u_j] n_j = \dot{f}_j n_j - [\delta_{kp}\delta_{lq} - \delta_{kq}\delta_{lp}] t_p^{(1)} t_q^{(2)} \partial_k \Psi_l = \dot{f}_j n_j + \{f_l, \Psi_l\} , \tag{A.33}$$

with Poisson brackets defined as:

$$\{f, g\} := \frac{\partial f}{\partial s_1}\frac{\partial g}{\partial s_2} - \frac{\partial f}{\partial s_2}\frac{\partial g}{\partial s_1} . \tag{A.34}$$

The classical form of the dynamic boundary condition:

$$T_{ij} n_j = \sigma_s \kappa n_i , \tag{A.35}$$

involving the stress tensor T_{ij}, surface tension σ_s and curvature κ, can be reformulated in terms of the tensor potential: substituting the term $-p\delta_{ij}$ in the stress tensor (3.40) by means of (3.45) and then replacing $\varepsilon_{jlk}\partial_l \varphi_k$ according to (3.59) by $-\partial_k \tilde{a}_{kj}$, the identity:

$$T_{ij} = \varrho u_i u_j + U\delta_{ij} - \varepsilon_{ilk}\varepsilon_{jpq}\partial_l \partial_p a_{kq} + \partial_i \left[\eta u_j + \partial_k \tilde{a}_{kj}\right] + \partial_j \left[\eta u_i + \partial_k \tilde{a}_{ki}\right] \tag{A.36}$$

results. Inserting (A.36) into (A.35), provides the general form of the dynamic boundary condition for unsteady flows.

In the case of steady flow, the kinematic boundary condition simplifies to: $0 = u_i n_i = \{f_l, \Psi_l\}$. Utilising this and (3.66), the dynamic boundary condition resulting from (A.35) and (A.36) takes the form:

$$\varepsilon_{ikl}\varepsilon_{jpq}\partial_k \partial_p a_{lq} n_j = (U - \sigma_s \kappa) n_i . \tag{A.37}$$

Next, the left-hand side of equation (A.37) can be written as:

$$\begin{aligned} \varepsilon_{ikl}\varepsilon_{jpq}\varepsilon_{jnm}\partial_k \partial_p a_{lq} t_n^{(1)} t_m^{(2)} &= \varepsilon_{ikl}\left[\delta_{pn}\delta_{qm} - \delta_{pm}\delta_{qn}\right]\partial_k \partial_p a_{lq} t_n^{(1)} t_m^{(2)} \\ = \varepsilon_{ikl}\left[t_m^{(2)}\frac{\partial}{\partial s_1}\left(\partial_k a_{lm}\right) - t_n^{(1)}\frac{\partial}{\partial s_2}\left(\partial_k a_{ln}\right)\right] &= \varepsilon_{ikl}\left[\frac{\partial}{\partial s_1}\left(\partial_k a_{lm} t_m^{(2)}\right) - \frac{\partial}{\partial s_2}\left(\partial_k a_{lm} t_m^{(1)}\right)\right] , \end{aligned}$$

in which the relationships:

$$t_m^{(\lambda)}\partial_m(\cdots) = \frac{\partial}{\partial s_\lambda}(\cdots) , \qquad \frac{\partial t_m^{(1)}}{\partial s_2} - \frac{\partial t_m^{(2)}}{\partial s_1} = 0 ,$$

have been used. Making use of the following relationship for the curvature:

$$-\kappa n_i = \frac{\partial^2 f_i}{\partial s_1^2} + \frac{\partial^2 f_i}{\partial s_2^2} = \frac{\partial t_i^{(1)}}{\partial s_1} + \frac{\partial t_i^{(2)}}{\partial s_2} = -\varepsilon_{ikl}\left[\frac{\partial}{\partial s_1}\left(n_k t_l^{(2)}\right) - \frac{\partial}{\partial s_2}\left(n_k t_l^{(1)}\right)\right],$$

together with the introduction of auxiliary functions $U_j(s_1, s_2)$ implicitly as solutions of the condition:

$$2U n_i = \varepsilon_{ijk}\left[\frac{\partial U_j}{\partial s_1} t_k^{(2)} - \frac{\partial U_j}{\partial s_2} t_k^{(1)}\right], \tag{A.38}$$

enables the dynamic boundary condition (A.37) to be written mathematically in the following integral form:

$$\frac{\partial}{\partial s_1}\left\{\varepsilon_{ikl}\left[\partial_k a_{lm} t_m^{(2)} + \left(\sigma_s n_k - \frac{U_k}{2}\right) t_l^{(2)}\right]\right\} - \frac{\partial}{\partial s_2}\left\{\varepsilon_{ikl}\left[\partial_k a_{lm} t_m^{(1)} + \left(\sigma_s n_k - \frac{U_k}{2}\right) t_l^{(1)}\right]\right\} = 0\,,$$

yielding:

$$\varepsilon_{ikl}\left[\partial_k a_{lm} t_m^{(\lambda)} + \left(\sigma_s n_k - \frac{U_k}{2}\right) t_l^{(\lambda)}\right] = \frac{\partial \chi_i}{\partial s_\lambda}\,, \qquad \lambda = 1, 2 \tag{A.39}$$

as the first integral of the dynamic boundary condition; containing the yet to be determined integration function $\chi_i = \chi_i(s_1, s_2)$. On applying the gauge transformation (3.53):

$$\begin{aligned} \varepsilon_{ikl}\partial_k a_{lm} t_m^{(\lambda)} \quad \rightarrow \quad & \varepsilon_{ikl}\partial_k a_{lm} t_m^{(\lambda)} + \varepsilon_{ikl}\partial_k\partial_l \alpha_m t_m^{(\lambda)} + t_m^{(\lambda)}\partial_m \varepsilon_{ikl}\partial_k \alpha_l \\ & = \varepsilon_{ikl}\partial_k a_{lm} t_m^{(\lambda)} + \frac{\partial}{\partial s_\lambda}\left(\varepsilon_{ikl}\partial_k \alpha_l\right), \end{aligned}$$

equation (A.39) becomes:

$$\varepsilon_{ikl}\left[\partial_k a_{lm} t_m^{(\lambda)} + \left(\sigma_s n_k - \frac{U_k}{2}\right) t_l^{(\lambda)}\right] = \frac{\partial}{\partial s_\lambda}\left[\chi_i - \varepsilon_{ikl}\partial_k \alpha_l\right], \qquad \lambda = 1, 2\,. \tag{A.40}$$

Via a proper choice of α_l, the right hand of equation (A.40) can be gauged to zero, leading to the simplified form:

$$\varepsilon_{ikl}\left[\partial_k a_{lm} t_m^{(\lambda)} + \left(\sigma_s n_k - \frac{U_k}{2}\right) t_l^{(\lambda)}\right] = 0\,, \qquad \lambda = 1, 2\,. \tag{A.41}$$

From a numerical standpoint the above formulation of the dynamic boundary condition contains an inconvenience, namely the necessity of having to construct the two tangent vectors $t_i^{(\lambda)}$. A more convenient and therefore more general formulation is obtained by contraction of (A.41) with $\mathrm{d}s_\lambda$, resulting in:

$$\varepsilon_{ikl}\left[\partial_k a_{lm}\mathrm{d}x_m + \left(\sigma_s n_k - \frac{U_k}{2}\right)\mathrm{d}x_l\right] = 0\,, \tag{A.42}$$

where the fact that $t_l^{(\lambda)} \mathrm{d}s_\lambda = \mathrm{d}x_l$ has been introduced. Note that the above parametrisation (A.28) of the free surface is not required for the general form (A.42) of the first integral of the dynamic boundary condition. For instance, the free surface may equally well be given in an explicit form, such as $x_3 = f(x_1, x_2)$, leading to $\mathrm{d}x_3 = (\partial f/\partial x_1)\,\mathrm{d}x_1 + (\partial f/\partial x_2)\,\mathrm{d}x_2$.

A.2.3 First integral with non-planar surfaces

(a) Lagrangian in terms of curvilinear coordinates

The velocity field (3.37) results in:

$$u_i = \varepsilon_{ijk}\partial_j \left[\psi\left(q_\mu\right)\partial_k q_3\right] = \varepsilon_{ijk}\partial_j\psi\left(q_\mu\right)\partial_k q_3 = \varepsilon_{ijk}\partial_j q_\mu \partial_k q_3 \partial_\mu \psi\,, \tag{A.43}$$

while $\partial_i \Phi = \partial_i q_\mu \partial_\mu \Phi$. Considering the definitions (3.128), (3.129), the useful relations:

$$\begin{aligned} u_i \partial_i q_\nu &= E_{\nu\mu}\partial_\mu \psi\,, \\ \partial_i \Phi \partial_i q_\nu &= G_{\nu\mu}\partial_\mu \Phi\,, \end{aligned}$$

are obtained from which the following identities are derived:

$$\begin{aligned} \bar{a}_{ij}u_i u_j &= a_{\mu\nu}u_i\partial_i q_\mu u_j \partial_j q_\nu = E_{\mu\lambda}E_{\nu\kappa}a_{\mu\nu}\partial_\lambda\psi\partial_\kappa\psi\,, \\ a_{\mu\nu}\partial_i q_\mu \partial_i \partial_j q_\nu &= \frac{1}{2}a_{\mu\nu}\partial_i q_\mu \partial_i \partial_j q_\nu + \frac{1}{2}a_{\mu\nu}\left[\partial_j\left(\partial_i q_\mu \partial_i q_\nu\right) - \partial_j \partial_i q_\mu \partial_i q_\nu\right] \\ &= \frac{1}{2}a_{\mu\nu}\underbrace{\left[\partial_i q_\mu \partial_i \partial_j q_\nu - \partial_i \partial_j q_\mu \partial_i q_\nu\right]}_{\text{skew-symmetric w.r.t. } \mu\leftrightarrow\nu} + \frac{1}{2}a_{\mu\nu}\partial_j q_\lambda \partial_\lambda G_{\mu\nu}\,, \\ \partial_i \bar{a}_{ij} &= \partial_i a_{\mu\nu}\partial_i q_\mu \partial_j q_\nu + a_{\mu\nu}\Delta q_\mu \partial_j q_\nu + a_{\mu\nu}\partial_i q_\mu \partial_i \partial_j q_\nu \\ &= \underbrace{\partial_i q_\mu \partial_i q_\lambda}_{G_{\mu\lambda}}\partial_\lambda a_{\mu\nu}\partial_j q_\nu + a_{\mu\nu}\left[\Delta q_\mu \partial_j q_\nu + \frac{1}{2}\partial_j q_\lambda \partial_\lambda G_{\mu\nu}\right]\,, \\ u_j \partial_i \bar{a}_{ij} &= G_{\mu\lambda}E_{\nu\kappa}\partial_\kappa\psi\partial_\lambda a_{\mu\nu} + \left[E_{\nu\kappa}\partial_\kappa\psi\Delta q_\mu + \frac{1}{2}E_{\lambda\kappa}\partial_\kappa\psi\partial_\lambda G_{\mu\nu}\right]a_{\mu\nu}\,, \\ \partial_j\Phi\partial_i\bar{a}_{ij} &= G_{\mu\lambda}G_{\nu\kappa}\partial_\kappa\Phi\partial_\lambda a_{\mu\nu} + \left[G_{\nu\kappa}\partial_\kappa\Phi\Delta q_\mu + \frac{1}{2}G_{\lambda\kappa}\partial_\kappa\Phi\partial_\lambda G_{\mu\nu}\right]a_{\mu\nu}\,. \end{aligned}$$

Next, consider the curl of the tensor potential:

$$\varepsilon_{ikl}\partial_k \bar{a}_{lm} = \varepsilon_{ikl}\partial_k q_\lambda \partial_\lambda \bar{a}_{lm}\,, \tag{A.44}$$

being aware that $\bar{a}_{lm}$ does not depend on q_3. It then follows that:

$$\partial_i q_\kappa \varepsilon_{ikl}\partial_k \bar{a}_{lm} = \overbrace{\varepsilon_{ikl}\partial_i q_\kappa \partial_k q_\lambda}^{\parallel \partial_l q_3}\overbrace{\partial_\lambda \bar{a}_{lm}}^{\perp \partial_l q_3} = 0\,, \tag{A.45}$$

$$\varepsilon_{ikl}\partial_k\bar{a}_{lm}\partial_m q_3 = \varepsilon_{ikl}\partial_k\underbrace{(\bar{a}_{lm}\partial_m q_3)}_{0} - \varepsilon_{ikl}a_{\mu\nu}\partial_l q_\mu\underbrace{\partial_m q_\nu\partial_k\partial_m q_3}_{\partial_k(\partial_m q_\nu\partial_m q_3)-\partial_m q_3\partial_m\partial_k q_\nu} = 0\,, \tag{A.46}$$

from which the form for the curl of the tensor potential:

$$\varepsilon_{ikl}\partial_k\bar{a}_{lm} = c_\mu\partial_i q_3\partial_m q_\mu\,,$$

is obtained with respective quantities c_μ. As a consequence of this, the quadratic term in the Lagrangian (3.74) vanishes according to:

$$\begin{aligned}\varepsilon_{ilk}\varepsilon_{jpq}\partial_l\bar{a}_{ij}\partial_p\bar{a}_{kq} =& -\varepsilon_{kli}\partial_l\bar{a}_{ij}\varepsilon_{jpq}\partial_p\bar{a}_{qk}\\ =& c_\mu\partial_k q_3\partial_j q_\mu c_\nu\partial_j q_3\partial_k q_\nu = c_\mu c_\nu\overbrace{\partial_j q_\mu\partial_j q_3}^{0}\overbrace{\partial_k q_\nu\partial_k q_3}^{0} = 0\,.\end{aligned}$$

Finally the Lagrangian (3.74) results in:

$$\ell = \varrho E_{\mu\lambda}E_{\nu\kappa}a_{\mu\nu}\partial_\lambda\psi\partial_\kappa\psi + \left[\eta E_{\nu\kappa}\partial_\kappa\psi - G_{\nu\kappa}\partial_\kappa\Phi\right]\left[2\left\{G_{\mu\lambda}\partial_\lambda + \Delta q_\mu\right\}a_{\mu\nu} + \partial_\nu G_{\mu\lambda}a_{\mu\lambda}\right]\,. \tag{A.47}$$

(b) First integral of the dynamic boundary condition in terms of spherical coordinates

Assuming the free surface to be given by $r = r_0 f(\vartheta)$ and $U = \varrho g r\cos\vartheta$ as the usual form of the potential energy density, the auxiliary quantities $\tilde{U}_\nu$ are determined via (3.134) as follows. First use is made of the identity:

$$\vec{t}^{(1)}\mathrm{d}s = r_0 f'(\vartheta)\vec{e}_r\mathrm{d}\vartheta + r_0 f(\vartheta)\vec{e}_\vartheta\mathrm{d}\vartheta = r_0\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left[f(\vartheta)\vec{e}_r\right]\mathrm{d}\vartheta\,,$$

and the decomposition $\vec{e}_r = \sin\vartheta\vec{e}_* + \cos\vartheta\vec{e}_z$ for the unit vector $\vec{e}_r$, where $\vec{e}_* := \cos\varphi\vec{e}_x + \sin\varphi\vec{e}_y$, delivering:

$$\begin{aligned}\frac{\vec{U}}{\varrho g r_0^2} &= \int f(\vartheta)\cos\vartheta\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left[f(\vartheta)\vec{e}_r\right]\mathrm{d}\vartheta\\ &= \int f(\vartheta)\cos\vartheta\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left[f(\vartheta)\sin\vartheta\right]\mathrm{d}\vartheta\,\vec{e}_* + \int f(\vartheta)\cos\vartheta\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left[f(\vartheta)\cos\vartheta\right]\mathrm{d}\vartheta\,\vec{e}_z\\ &= \underbrace{\int f(\vartheta)f'(\vartheta)\sin\vartheta\cos\vartheta\mathrm{d}\vartheta}_{\frac{f^2}{2}\sin\vartheta\cos\vartheta-\int\frac{f^2}{2}[\cos^2\vartheta-\sin^2\vartheta]\mathrm{d}\vartheta}\vec{e}_* + \int f(\vartheta)^2\cos^2\vartheta\mathrm{d}\vartheta\,\vec{e}_* + \frac{1}{2}\left[f(\vartheta)\cos\vartheta\right]^2\vec{e}_z\\ &= \frac{1}{2}f(\vartheta)^2\cos\vartheta\underbrace{\left[\sin\vartheta\vec{e}_* + \cos\vartheta\vec{e}_z\right]}_{\vec{e}_r} + \frac{1}{2}\int f(\vartheta)^2\mathrm{d}\vartheta\,\vec{e}_*\,.\end{aligned}$$

Decomposition of the above vector according to $\vec{U} = \tilde{U}_r\nabla r + \tilde{U}_\vartheta\nabla\vartheta = \tilde{U}_r\vec{e}_r + r^{-1}\tilde{U}_\vartheta\vec{e}_\vartheta$ yields:

$$\frac{\tilde{U}_r}{\varrho g r_0^2} = \vec{U}\cdot\vec{e}_r = \frac{1}{2}f(\vartheta)^2\cos\vartheta + \frac{\sin\vartheta}{2}\int f(\vartheta)^2\mathrm{d}\vartheta\,,$$

$$\frac{\tilde{U}_\vartheta}{\varrho g r_0^2} = r_0 f(\vartheta)\vec{U}\cdot\vec{e}_\vartheta = \frac{r_0 f(\vartheta)\cos\vartheta}{2}\int f(\vartheta)^2 \mathrm{d}\vartheta\,,$$

while the normal vector reads:

$$\vec{n} = \frac{\vec{e}_r - \frac{f'(\vartheta)}{f(\vartheta)}\vec{e}_\vartheta}{\sqrt{1+\frac{f'(\vartheta)^2}{f(\vartheta)^2}}} = \frac{f(\vartheta)\vec{e}_r - f'(\vartheta)\vec{e}_\vartheta}{\sqrt{f(\vartheta)^2+f'(\vartheta)^2}} = \frac{f(\vartheta)\nabla r - r_0 f(\vartheta)f'(\vartheta)\nabla\vartheta}{\sqrt{f(\vartheta)^2+f'(\vartheta)^2}}\,.$$

Finally, the dynamic boundary condition (3.133) result in:

$$2\frac{\partial\Phi}{\partial r} + \frac{\sigma f(\vartheta)}{\sqrt{f(\vartheta)^2+f'(\vartheta)^2}} - \frac{\varrho g r_0^2}{2}\left[f(\vartheta)^2\cos\vartheta + \int f(\vartheta)^2\mathrm{d}\vartheta\sin\vartheta\right] = 0\,, \tag{A.48}$$

$$2\frac{\partial\Phi}{\partial\vartheta} - \frac{\sigma r_0 f(\vartheta)f'(\vartheta)}{\sqrt{f(\vartheta)^2+f'(\vartheta)^2}} - \frac{\varrho g r_0^3}{2}f(\vartheta)\cos\vartheta\int f(\vartheta)^2\mathrm{d}\vartheta = 0\,. \tag{A.49}$$

(c) Axially symmetric flows in terms of cylindrical coordinates

By making the choice $q_1 = z$, $q_2 = r$ with $q_3 = \varphi$ as neutral coordinate, $G_{\mu\nu} = \delta_{\mu\nu}$ and $E_{\mu\nu} = r^{-1}\varepsilon_{\mu\nu}$ are obtained. Considering $\Delta z = 0$ and $\Delta r = 1/r$, the corresponding Lagrangian (3.130) takes the form:

$$\begin{aligned}\ell = {}& \varrho\left[a_{11}\left(u_z^2 - u_r^2\right) + 2a_{12}u_r u_z\right] \\ & + 2\left[\eta u_z - \partial_z\Phi\right]\left[\partial_z a_{11} + \frac{1}{r}\partial_r\left(r a_{12}\right)\right] + 2\left[\eta u_r - \partial_r\Phi\right]\left[\partial_z a_{12} - \frac{1}{r}\partial_r\left(r a_{11}\right)\right]\,,\end{aligned}$$

with the velocities given by:

$$u_z = \frac{1}{r}\frac{\partial\psi}{\partial r}\,, \qquad\qquad u_r = -\frac{1}{r}\frac{\partial\psi}{\partial z}\,. \tag{A.50}$$

Variation of the action integral $I = 2\pi\iint_A \ell\, r\mathrm{d}r\mathrm{d}z$ with respect to a_{1i} delivers, after division by $r/2$:

$$\delta a_{11}: \frac{\varrho}{2}\left(u_z^2 - u_r^2\right) = \partial_z\left[\eta u_z - \partial_z\Phi\right] - \partial_r\left[\eta u_r - \partial_r\Phi\right]\,, \tag{A.51}$$

$$\delta a_{12}: \varrho u_r u_z = \partial_r\left[\eta u_z - \partial_z\Phi\right] + \partial_z\left[\eta u_r - \partial_r\Phi\right]\,. \tag{A.52}$$

A.2.4 Finite difference scheme

A classical velocity-pressure staggered-grid scheme for the finite difference solution of viscous flow problems [95] is extended to encompass the tensor entries a_{ij} in a consistent way which is partly inspired by a common numerical method for solving the velocity-stress-formulation of wave propagation through elastic media, as for example utilised by Graves [110]. As well as the stabilising effect inherent with the use of a staggered grid

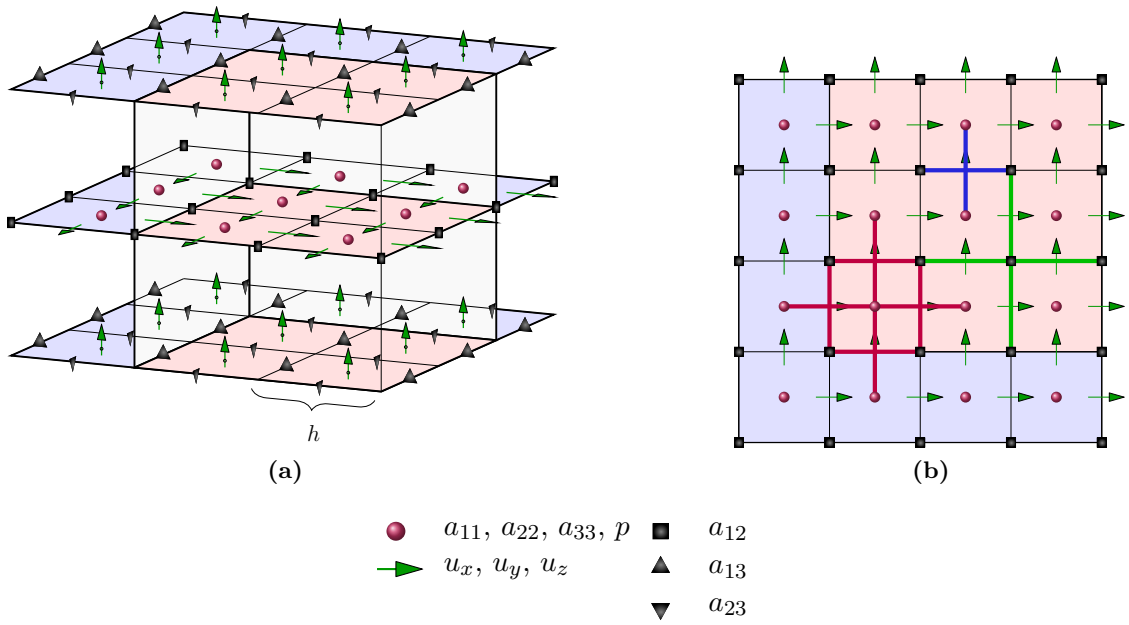


Figure A.1: Schematic of the 3D staggered grid arrangement used to solve the lid-driven cavity problem. (a) Shows three contiguous grids with the pressure and the diagonal tensor entries located at, and identified by, the red spheres, velocities at the sites indicated by green arrows, the off-diagonal tensor entries at the corresponding black squares and triangles. The solution domain of interest is shaded red while the associated boundary region containing the necessary ghost points is shaded blue. (b) View of just one xy-plane, with the dependencies of selected finite difference stencils highlighted at different grid points for the sake of clarity: the stencil for Eq. (3.112) (for a_{12}) is shown in green, that for Eq. (3.113) in red and that for Eq. (3.115) (for u_2) in blue.

arrangement, the method also economises on the number of unknowns in contrast to the use of any alternative non-staggered grid scheme. The resulting 3D grid arrangement, see Fig. A.1, is such that the diagonal tensor components and the pressure are discretised at identical cell centred grid points (i, j, k), the velocities at face centred grid points $(i+1, j, k)$, $(i, j+1, k)$, $(i, j, k+1)$ and the off-diagonal tensor components at cell edges $(i+1, j+1, k)$, $(i, j+1, k+1)$, $(i+1, j, k+1)$.

In contrast to classical discretisation of the vector-valued NS equations, in the present case the "mapping" between equations and unknown fields is less obvious. Here, the six equations of (3.112) are discretised at the grid points of the corresponding tensor potential entries, equation (3.113) at the pressure grid points and equations (3.114) at the boundary velocity grid points only; correspondingly, recovery of the velocities, equation (3.115), is also performed at the velocity grid points. Boundary conditions are incorporated via an appropriate ghost cell method involving two grid levels of additional boundary points where conditions (3.84)-(3.86) are similarly specified. Figure A.1a highlights a section

of the solution domain, shaded red, with the associated boundary region, shaded blue, illustrating the staggered grid arrangement in terms of the various grid point locations.

With reference to the above defined correlation between equations and unknowns, equations (83) to (85) are discretised at different subgrids as indicated in Fig. A.1a and all occurring derivatives are approximated by second order central difference stencils in which the nearest available grid points are utilised; the staggered grid arrangement accounts for varying step lengths for the diverse stencils, as illustrated in Figure A.1b. The velocities u_i, used iteratively in equation (3.112), are calculated from the tensor potential entries subsequent to each iteration step, which is achieved by application of equation (3.115) at all points indicated by a green arrow. The resulting velocity field is then interpolated onto the remaining grid points by a simple weighting of neighbouring points.

The finite difference analogues of the system (3.112)-(3.115) can be written in a compact way by defining a number of discrete operators for an arbitrary 3D scalar function $f : \Omega \to \mathbb{R}$; these are valid point-wise for a given set of indices $[i, j, k]$ belonging to grid coordinates (x_i, y_j, z_k). Note, that in the finite difference description provided a grid with uniform step length h in all three coordinate directions is assumed for convenience only (see Fig. A.1); generalisation to a more complex grid pattern follows in a straightforward manner. The standard second order central difference operators for the first and second order partial derivatives of f are given by:

$$\begin{aligned}
\partial_{h,1}[i,j,k]f &:= \frac{1}{h}\left[f(x_i + h/2, y_j, z_k) - f(x_i - h/2, y_j, z_k)\right], \\
\partial_{h,2}[i,j,k]f &:= \frac{1}{h}\left[f(x_i, y_j + h/2, z_k) - f(x_i, y_j - h/2, z_k)\right], \\
\partial_{h,3}[i,j,k]f &:= \frac{1}{h}\left[f(x_i, y_j, z_k + h/2) - f(x_i, y_j, z_k - h/2)\right], \\
\partial_{h,11}[i,j,k]f &:= \frac{1}{h^2}\left[f(x_i + h, y_j, z_k) - 2f(x_i, y_j, z_k) + f(x_i - h, y_j, z_k)\right], \\
\partial_{h,22}[i,j,k]f &:= \frac{1}{h^2}\left[f(x_i, y_j + h, z_k) - 2f(x_i, y_j, z_k) + f(x_i, y_j - h, z_k)\right], \\
\partial_{h,33}[i,j,k]f &:= \frac{1}{h^2}\left[f(x_i, y_j, z_k + h) - 2f(x_i, y_j, z_k) + f(x_i, y_j, z_k - h)\right],
\end{aligned}$$

which allows the discrete Laplacian to be written as:

$$\Delta_h[i,j,k]f := \sum_{l=1}^{3} \partial_{h,ll}[i,j,k]f\,. \tag{A.53}$$

Second order discretisation of the mixed derivatives is performed in the standard way, giving:

$$\begin{aligned}
\partial_{h,12}[i,j,k]f := \frac{1}{h^2}\big[&f(x_i + h/2, y_j + h/2, z_k) - f(x_i + h/2, y_j - h/2, z_k) \\
&- f(x_i - h/2, y_j + h/2, z_k) + f(x_i - h/2, y_j - h/2, z_k)\big],
\end{aligned}$$

$$\partial_{h,13}[i,j,k]f := \frac{1}{h^2}\big[f(x_i+h/2,y_j,z_k+h/2) - f(x_i+h/2,y_j,z_k-h/2) \\ -f(x_i-h/2,y_j,z_k+h/2) + f(x_i-h/2,y_j,z_k-h/2)\big]\,,$$

$$\partial_{h,23}[i,j,k]f := \frac{1}{h^2}\big[f(x_i,y_j+h/2,z_k+h/2) - f(x_i,y_j+h/2,z_k-h/2) \\ -f(x_i,y_j-h/2,z_k+h/2) + f(x_i,y_j-h/2,z_k-h/2)\big]\,.$$

In addition to the above well-known finite difference stencils the following interpolation operators for functions and their first order derivatives are introduced for convenience:

$$I_{h,1}[i,j,k]f := \frac{1}{2}\big[f(x_i+h/2,y_j,z_k) + f(x_i-h/2,y_j,z_k)\big]\,, \tag{A.54}$$

$$I_{h,2}[i,j,k]f := \frac{1}{2}\big[f(x_i,y_j+h/2,z_k) + f(x_i,y_j-h/2,z_k)\big]\,, \tag{A.55}$$

$$I_{h,3}[i,j,k]f := \frac{1}{2}\big[f(x_i,y_j,z_k+h/2) + f(x_i,y_j,z_k-h/2)\big]\,, \tag{A.56}$$

$$J_h^{r,s}[i,j,k]f := I_{h,r}[i,j,k]\partial_{h,s}f\,; \tag{A.57}$$

with (A.57) in particular facilitating a very compact discrete form of the tensor-valued equation (3.112), i.e., for $\alpha,\beta = 1,2,3$ it is:

$$\Delta_h[\gamma]a_{\alpha\beta}^{(n+1)} - \mathrm{Re}\left[I_{h,\beta}[\gamma]u_\alpha^{(n)}\sum_{l=1}^{3}J_h^{\alpha,l}[\gamma]a_{\beta l}^{(n+1)} + I_{h,\alpha}[\gamma]u_\beta^{(n)}\sum_{l=1}^{3}J_h^{\beta,l}[\gamma]a_{\alpha l}^{(n+1)}\right] \\ + \Big[p^{(n+1)}(x_i,y_j,z_k) + U(x_i,y_j,z_k)\Big]\,\delta_{\alpha\beta} = \mathrm{Re}\,I_{h,\beta}[\gamma]u_\alpha^{(n)}I_{h,\alpha}[\gamma]u_\beta^{(n)}\,, \tag{A.58}$$

with the abbreviation $\gamma = (i,j,k)$ used for a given index set. Recall, that equations (A.58) are not formed at all grid points (i,j,k), but rather at the respective subsets of grid points belonging to $a_{\alpha\beta}$ according to Fig. A.1, whereas the discrete form of equation (3.113):

$$\sum_{k,l=1}^{3}\partial_{h,kl}[\gamma]a_{kl}^{(n+1)} = 0\,,$$

is collocated at the pressure grid points only. Finally, the discrete form of the vector-valued equation (3.115), similar to that of (3.114), is given by:

$$u_\alpha^{(n+1)} = -\sum_{l=1}^{3}\partial_{h,l}[\gamma]a_{\alpha l}^{(n+1)}\,,$$

for $\alpha = 1,2,3$ and collocated at the u_α velocity grid points.

B Functional analysis and elliptic PDE theory

B.1 Brief compendium of functional analysis

For analysis of the least-squares FEM in Chapter 4 a fundamental acquaintance with the field of functional analysis is required; this is also essential for the more advanced considerations relating to the Agmon-Douglis-Nirenberg theory in B.2 and the regularity investigations in B.3 and B.4. The following compendium is mainly based on the standard textbooks by Hackbusch [119] and Braess [44].

B.1.1 Banach and Hilbert spaces

Definition B.1. (Norm) *Given a linear space (vector space) V over the field $\mathbb{K}$ of real or complex numbers, a* norm *on V is a mapping $\|\cdot\| : V \to \mathbb{R}_0^+, v \mapsto \|v\|$ such that for all $u, v \in V$ and $\lambda \in \mathbb{K}$:*

$$\begin{aligned} \|v\| = 0 \quad &\Rightarrow \quad v = 0 && (\textit{definiteness}) \\ \|\lambda v\| &= |\lambda| \cdot \|v\| && (\textit{absolute homogeneity}) \\ \|u + v\| &\leq \|u\| + \|v\| && (\textit{subadditivity}) \end{aligned}$$

Definition B.2. (Banach space) *A sequence $\{v_n \in V : n \geq 1\}$ is called a* Cauchy sequence *if:*

$$\sup \{\|v_n - v_m\|_V : n, m \geq k\} \to 0 \quad \textit{for} \quad k \to \infty \,.$$

A space V is said to be complete*, if every Cauchy sequence converges to some $v_\infty \in V$. A* Banach space *is a complete and normed space.*

Definition B.3. (Inner product) *Given a linear space V over the field $\mathbb{K}$ of real or complex numbers, the mapping $(\cdot, \cdot) : V \times V \to \mathbb{K}$ defines an* inner product *on V, if for all $u, v, w \in V$ and $\lambda \in \mathbb{K}$:*

$$(v, v) \geq 0 \quad \textit{and} \quad (v, v) = 0 \Leftrightarrow v = 0 \qquad (\textit{positive definiteness})$$

$$(\lambda u + v, w) = \lambda(u, w) + (v, w) \qquad (\textit{linearity})$$
$$(u, v) = \overline{(v, u)} \qquad (\textit{conjugate symmetry})$$

Definition B.4. (Hilbert space) *A Banach space V is named a* Hilbert space*, if an inner product $(\cdot,\cdot)_V$ exists on V, such that a norm is induced by $\|v\|_V = \sqrt{(v,v)_V}$ for all $v \in V$.*

It is easily seen that in the latter case the *inequality of Cauchy-Schwarz* is valid:

$$|(u, v)_V| \leq \|u\|_V \|v\|_V \quad \text{for all } u, v \in V\,.$$

B.1.2 Sobolev spaces

The so-called *Sobolev* function spaces, which play an essential role in the solution theory of PDEs, are based on the Lebesgue spaces L^p of Lebesgue-integrable functions. In the following let Ω always be an open subset of $\mathbb{R}^n$ with a piecewise smooth boundary. The space $L^p(\Omega)$ comprises all Lebesgue-measurable functions f with a finite integral over $|f|^p$, i.e.:

$$L^p(\Omega, \mathcal{A}, \mu) := \left\{ f : \Omega \to \mathbb{K}, \int_\Omega |f(x)|^p \,\mathrm{d}\mu(x) < \infty \right\},$$

in which the triple $(\Omega, \mathcal{A}, \mu)$ forms a measurable space involving a σ-algebra $\mathcal{A}$ and a measure μ, and $\mathbb{K}$ is the space of real or complex numbers. Basically, two functions $u, v \in L^p(\Omega)$ are considered identical ($u = v$), if $u(x) = v(x)$ holds for nearly all $x \in \Omega$ except for a set with a Lebesgue measure of zero; an identification that guarantees the mapping:

$$\begin{aligned} \|\cdot\|_{L^p} : L^p &\to \mathbb{R} \\ f &\mapsto \left(\int_\Omega |f(x)|^p \,\mathrm{d}\mu(x) \right)^{1/p}, \end{aligned}$$

defines a norm on L^p for all $p \geq 1$, rather than a semi-norm. In fact, for $1 \leq p \leq \infty$, L^p forms a Banach space but only for the case $p = 2$ can a corresponding inner product be defined giving the L^2-space crucial importance:

Proposition B.5. *$L^2(\Omega)$ forms a Hilbert space equipped with the inner product:*

$$(u, v)_{L^2(\Omega)} := \int_\Omega u(x)\overline{v(x)} \,\mathrm{d}x \quad \textit{for all } u, v \in L^2(\Omega)\,, \tag{B.3}$$

and the associated norm:

$$\|u\|_{L^2(\Omega)} := \sqrt{\int_\Omega |u(x)|^2 \,\mathrm{d}x}\,.$$

Some parts of the analysis in Chapter 5 involve products of L^p-functions which need not lie in the same space, instead the following extended *Hölder inequality* is valid:

Proposition B.6. (Hölder inequality) *Assume that $f_1, f_2, \dots f_k$ are functions such that:*

$$f_i \in L^{p_i}\,,\ 1 \le i \le k \text{ with } \frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_k} \le 1\,; \tag{B.4}$$

then the product $f = f_1 f_2 \cdots f_k$ belongs to L^p and:

$$\|f\|_p \le \|f_1\|_{p_1} \|f_2\|_{p_2} \cdots \|f_k\|_{p_k}\,. \tag{B.5}$$

In particular, if $f \in L^p \cap L^q$ with $1 \le p \le q \le \infty$, then $f \in L^r$ for all $r, p \le r \le q$, and the following interpolation inequality holds:

$$\|f\|_r \le \|f\|_p^{\alpha} \|f\|_q^{1-\alpha}\,,\ \text{where } \frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q}\,,\ 0 \le \alpha \le 1\,.$$

Essential for the definition of Sobolev spaces is the concept of *weak* derivatives:

Definition B.7. (Multi-indices) *A vector of indices $\alpha = (\alpha_1, \dots \alpha_n)$ with non-negative components α_i is said to be a multi-index of order:*

$$|\alpha| = \alpha_1 + \cdots + \alpha_n\,.$$

For a given multi-index α a partial derivative of order $|\alpha|$ is defined according to:

$$D^{\alpha} u(x) = \frac{\partial^{|\alpha|} u(x)}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}\,.$$

Definition B.8. (Weak derivative) *$u \in L^2(\Omega)$ is said to have a* weak derivative *$v := D^{\alpha} u \in L^2(\Omega)$, if v fulfils:*

$$(w, v)_{L^2(\Omega)} = (-1)^{|\alpha|} (D^{\alpha} w, u)_{L^2(\Omega)} \quad \text{for all } w \in C_0^{\infty}(\Omega)\,.$$

In the above definition $C_0^{\infty}(\Omega)$ denotes the function space of infinitely often continuously differentiable functions with compact support, that is:

$$\begin{aligned} C_0^{\infty}(\Omega) &:= \{u \in C^{\infty}(\Omega) : \operatorname{supp}(u) \text{ compact}, \operatorname{supp}(u) \subset\subset \Omega\}\,, \\ \operatorname{supp}(u) &:= \overline{\{x \in \Omega : u(x) \neq 0\}}\,, \end{aligned}$$

in which the double inclusion $\Omega' \subset\subset \Omega$ is meant to say that Ω' is in the *inner* part of Ω, implying $\Omega' \subset \Omega$ and $\Omega' \cap \partial\Omega = \emptyset$. In a further step, spaces of weakly differentiable functions can be defined as follows:

Definition B.9. (Sobolev spaces) *Let $k \in \mathbb{N}_0$. Define $H^k(\Omega) \subset L^2(\Omega)$ as the set of functions with existing weak derivatives $D^\alpha u \in L^2(\Omega)$ for all $|\alpha| \leq k$:*

$$H^k(\Omega) = \{u \in L^2(\Omega) : D^\alpha u \in L^2(\Omega) \,\forall\, |\alpha| \leq k\}\,. \tag{B.6}$$

A similar definition to the above is possible based on any $L^p(\Omega)$-space in which case it is usually denoted as $W^{k,p}$, so that $H^k(\Omega) = W^{k,2}(\Omega)$. The above function spaces defined via weak derivatives are referred to as *Sobolev-spaces* and, as already suggested by the capital H, also form Hilbert spaces when equipped with suitable inner products and norms:

Proposition B.10. *$H^k(\Omega)$ forms a Hilbert space equipped with the inner product:*

$$(u,v)_k = (u,v)_{H^k(\Omega)} := \sum_{|\alpha|\leq k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}\,, \tag{B.7}$$

and the associated (Sobolev-)norm:

$$\|u\|_k := \|u\|_{H^k(\Omega)} := \sqrt{\sum_{|\alpha|\leq k} \|D^\alpha u\|^2_{L^2(\Omega)}}\,. \tag{B.8}$$

As the H^k-spaces are due to their identification as Hilbert spaces, see Definition B.4, complete, an alternative but insightful definition of them can be provided via *completion*. Let $X := \{u \in C^\infty(\Omega) : |u|_k < \infty\}$, then the completion of $X = C^\infty(\Omega) \cap H^k(\Omega)$ in $L^2(\Omega)$ with respect to the norm (B.8) is exactly $H^k(\Omega)$. Accordingly, the completion of $C_0^\infty(\Omega)$ in $L^2(\Omega)$ with respect to (B.8) is denoted as $H_0^k(\Omega)$. The latter one is especially important in the context of boundary value problems with Dirichlet-boundaries. The completed space $H_0^k(\Omega)$ is a subspace of $H^k(\Omega)$ and again forms a Hilbert space with the same inner product (B.7) and the norm (B.8). Particularly in the case of $k = 0$ it is:

$$H_0^0(\Omega) = H^0(\Omega) = L^2(\Omega)\,,$$

and the shorter notation of both choices $(u,v)_{L^2(\Omega)} = (u,v)_0$ and $\|u\|_{L^2(\Omega)} = \|u\|_0$ is used frequently. Moreover, by:

$$|u|_k = |u|_{H^k(\Omega)} := \sqrt{\sum_{|\alpha|=k} \|D^\alpha u\|^2_{L^2(\Omega)}}\,, \tag{B.9}$$

a semi-norm is defined and by applying the well-known Poincaré-Friedrich inequality it can be demonstrated that $|\cdot|_{H^k(\Omega)}$ and $\|\cdot\|_{H^k(\Omega)}$ represent equivalent norms on $H_0^k(\Omega)$, if the domain Ω is bounded:

Proposition B.11. (Poincaré-Friedrich) *Consider a bounded domain $\Omega \subset \mathbb{R}^n$, then the inequality:*

$$\|v\|_0 \leq s|v|_1 \quad \textit{for all } v \in H_0^1(\Omega)\,, \tag{B.10}$$

holds with $s = |\Omega|^{1/n}$ and the norms $|\cdot|_k$ and $\|\cdot\|_k$ are equivalent due to:

$$|v|_m \leq \|v\|_m \leq (1+s)^m |v|_m \quad \textit{for all } v \in H_0^m(\Omega)\,.$$

For clarification different connections are summarised in the subsequent complex:

$$\begin{array}{ccccccc} L^2(\Omega) = H^0(\Omega) & \supset & H^1(\Omega) & \supset & H^2(\Omega) & \supset & \dots \\ \| & & \cup & & \cup & & \\ H_0^0(\Omega) & \supset & H_0^1(\Omega) & \supset & H_0^2(\Omega) & \supset & \dots \end{array}$$

In a next step the previously defined Sobolev spaces H^k with $k \in \mathbb{N}_0$ are extended to non-integer indices:

Definition B.12. *Let $\Omega \subset \mathbb{R}^n$ and split $s \in \mathbb{R}_0^+$ into two parts $s = k + \lambda$ with $k \in \mathbb{N}_0$ and $0 < \lambda < 1$. Then an inner product and a corresponding norm, the so-called* Sobolev-Slobodeckij norm*, are defined by:*

$$(u,v)_s := \sum_{\alpha \leq k} \left[\int_\Omega D^\alpha u(x) D^\alpha v(x)\,\mathrm{d}x + \iint_{\Omega\times\Omega} \frac{[D^\alpha u(x) - D^\alpha u(y)][D^\alpha v(x) - D^\alpha v(y)]}{|x-y|^{n+2\lambda}}\,\mathrm{d}x\mathrm{d}y \right],$$
$$\|u\|_s := \|u\|_{H^s(\Omega)} := \sqrt{(u,u)_s}\,. \tag{B.11}$$

Obviously for the above definition only the existence of weak derivatives up to order k is necessary as was the case for the integer order Sobolev spaces. Thus, a fractional order Sobolev space H^s can be defined by the set of functions (B.6) equipped with the above inner product and the Sobolev-Slobodeckij norm which again forms a Hilbert space. For $s \in \mathbb{N}_0$ this definition coincides with Definition B.9 and Proposition B.10. Important properties of the fractional order Sobolev spaces are summarised in the following Propostion:

Proposition B.13. *Let $s \geq 0$.*

(i) $\{u \in C^\infty(\Omega) : \mathrm{supp}(u)$ *compact*, $|u|_s < \infty\}$ *is dense in $H^s(\Omega)$.*

(ii) C_0^∞ *is dense in $H_0^s(\Omega)$.*

(iii) $aD^\alpha(bu) \in H^{s-|\alpha|}(\Omega)$, *if* $|\alpha| \leq s$, $u \in H^s(\Omega)$, $a \in C^{t-|\alpha|}(\overline{\Omega})$, $b \in C^t(\overline{\Omega})$ *with either* $t = s \in \mathbb{N}_0$ *or* $t > s$.

(iv) $H^s(\Omega) \subset H^t(\Omega)$, $H_0^s(\Omega) \subset H_0^t(\Omega)$ *for* $s \geq t$.

Often the following interpolation result is useful when dealing with H_0^m-spaces:

Proposition B.14. *Let $m, k \in \mathbb{R}$. For $m \geq 1$ constants $c = c(m)$ exist, such that:*

$$|u|_k \leq c|u|_m^{k/m}|u|_0^{(m-k)/m} \quad \text{for all } 0 \leq k \leq m,\ u \in H_0^m(\Omega)\,. \tag{B.12}$$

It is important to know, how different spaces like Lebesgue, Sobolev and Hölder spaces are related to each other. In this context a key result is the following version of the *Rellich-Kondrachov embedding theorem* according to Evans [92]:

Theorem B.15. (Rellich-Kondrachov) *Let $\Omega \subset \mathbb{R}^n$ be a Lipschitz domain, $k \in \mathbb{Z}_+$ and $p \in [1, \infty)$ then the following embeddings are valid:*

(i) If $kp < n$, $W^{k,p}(\Omega) \subset L^q(\Omega)$ is a compact embedding for all $q \in [1, p^*)$, where $\frac{1}{p^*} = \frac{1}{p} - \frac{k}{n}$. For $q = p^*$ the embedding is continuous only.

(ii) If $kp = n$, $W^{k,p}(\Omega) \subset L^q(\Omega)$ is a compact embedding for all $q \in [p, \infty)$.

(iii) If $kp > n$, $W^{k,p}(\Omega) \subset C^m(\overline{\Omega})$ for every integer m that satisfies $0 \leq m < k - \frac{n}{p}$.

Although the Rellich-Kondrachov theorem is more often cited under the stronger condition of a C^1 domain boundary due to a simpler proof, see for instance Brezis [53], the above cited version is more helpful, particularly in the context of FEs where most often piecewise smoothly bounded domains occur which are generally in C^0 only. Also, the theorem is frequently reduced to the case (i), probably due to historical reasons, whereas the treatment of two-dimensional elliptic equations rather requires (ii), as seen in Sec. 4.4. Note, (iii) reveals a well-known fact as a special case, namely that $H^1(\Omega)$-functions need to be continuous only in one dimension.

B.1.3 Traces and continuation

The nature of boundary value problems requires a reasonable determination of boundary values $u|_{\partial\Omega}$ in the sense that the restriction of a function u in Ω to the boundary $\partial\Omega$, the so-called *trace*, results in a well-defined object. It is easy to see that a Hölder-continuous function $u \in C^s(\overline{\Omega})$ has a restriction $u|_{\partial\Omega} \in C^s(\partial\Omega)$ if $\partial\Omega$ is simply sufficiently smooth. On the other hand from $u \in H^s(\Omega)$ it does not automatically follow that $u|_{\partial\Omega}$; the identity $u = v$ in $H^s(\Omega)$ only requires that $u(x) = v(x)$ is equal nearly everywhere in Ω apart from a set of zero measure, as for instance the boundary $\partial\Omega$ is. Therefore $u(x) \neq v(x)$ can in principle hold for any $x \in \partial\Omega$. Neither can the boundary value $u(x)$, $x \in \partial\Omega$ be defined by continuous extension, as for instance $u \in H^1(\Omega)$ need not to be continuous. The inverse problem to the definition of $u|_{\partial\Omega}$ is the continuation: does for a given boundary value ϕ on $\partial\Omega$ a function $u \in H^s(\Omega)$ exist such that $\phi = u|_{\partial\Omega}$? In the negative case there is also no solution $u \in H^s(\Omega)$ to the Dirichlet boundary value problem.

Lemma B.16. *Given a bounded domain $\Omega \in C^t$ $[\omega \in C^{k,1}]$, the corresponding quantities $N \in \mathbb{N}$, U^i $(0 \le i \le N)$, U_i, α_i $(1 \le i \le N)$ exist with:*

(i) U^i *open, bounded* $(0 \le i \le N)$, $\quad \bigcup_{i=0}^N U^i \supset \overline{\Omega}$, $\quad U^0 \subset\subset \Omega$,

(ii) $U_i := U^i \cap \partial\Omega$ $(1 \le i \le N)$, $\quad \bigcup_{i=0}^N U_i = \partial\Omega$,

(iii) $\alpha_i : U_i \to \alpha_i(U_i) \subset \mathbb{R}^{n-1}$ *bijective for all* $i = 1, \dots, N$,

(iv) $\alpha_i \circ \alpha_j^{-1} \in C^t(\overline{\alpha_j(U_i \cap U_j)})$ *or likewise* $\alpha_i \circ \alpha_j^{-1} \in C^{k,1}(\overline{\alpha_j(U_i \cap U_j)})$.

A set of pairings $\{(U_i, \alpha_i) : 1 \le i \le N\}$ with the above properties is called a C^t or $C^{k,1}$ coordinate system of $\partial\Omega$.

Lemma B.17. (Partition of unity) *Given $\{U^i : 0 \le i \le N\}$ with the property (i) of Lemma B.16, then functions $\sigma_i \in C_0^\infty(\mathbb{R}^n)$, $0 \le N$ exist with:*

$$\operatorname{supp}(\sigma_i) \subset U^i\,, \quad \sum_{i=0}^N \sigma_i^2(x) = 1 \;\; \text{for all } x \in \overline{\Omega}\,. \tag{B.13}$$

The general construction of the σ_i can be found in [119]. A function u on $\partial\omega$ can be written in the form $\sum \sigma_i^2 u$ and every summand $\sigma_i^2 u$ is parametrisable over $\alpha_i(U_i) \subset \mathbb{R}^{n-1}$: $(\sigma_i^2) \circ \alpha_i^{-1} : \alpha_i(U_i) \subset\subset \mathbb{R}^{n-1} \to \mathbb{R}$. This gives rise to the following definition:

Definition B.18. *Let $\Omega \in C^t\,[\in C^{k,1}]$. (U_i, α_i) and σ_i fulfil the conditions of Lemma B.16 (ii)-(iv) and* (B.13). *Let $s \le t \in \mathbb{N}$ $[s \le k+1]$ or $s < t \notin \mathbb{N}$, $t > 1$. The Sobolev space $H^s(\partial\Omega)$ is the set of all functions $u : \partial\Omega \to \mathbb{R}$ such that $(\sigma_i u) \circ \alpha_i^{-1} \in H_0^s(\mathbb{R}^{n-1})$, $i \le i \le N$.*

Proposition B.19. *$H^s(\partial\Omega)$ defines a Hilbert space with the corresponding inner product:*

$$(u, v)_s := (u, v)_{H^s(\partial\Omega)} := \sum_{i=1}^N \left((\sigma_i u) \circ \alpha_i^{-1}, (\sigma_i v) \circ \alpha_i^{-1} \right)_{H^s(\mathbb{R}^{n-1})}\,.$$

If $\{(\tilde{U}_i, \tilde{\alpha}_i) : 1 \le i \le N\}$ is another C^t $[C^{k,1}]$ coordinate system of $\partial\Omega$ and $\{\tilde{\sigma}_i\}$ another partition of unity, the hereby defined space $\tilde{H}^s(\partial\Omega)$ is quantitatively equal to $H^s(\partial\Omega)$. The norms of $H^s(\partial\Omega)$ and $\tilde{H}^s(\partial\Omega)$ are equivalent.

A few statements about traces and continuation can be made in sufficiently smoothly bounded domains; here, γ denotes the trace operator restricting a function to its boundary: $\gamma u = u|_{\partial\Omega}$:

Proposition B.20. (Trace inequality) *Let $\Omega \in C^t$ with $1/2 < s < t \in \mathbb{N}$ or $1/2 < s < t$ [alternatively $\Omega \in C^{k,1}, 1/2 < s = k+1 \in \mathbb{N}$]. The following statements hold:*

(i) The restriction of $u \in H^s(\Omega)$ to the boundary, i.e. γu, belongs to $H^{s-1/2}(\partial\Omega)$; thus the restriction mapping is of the form:

$$\begin{aligned} \gamma : H^s(\Omega) &\to H^{s-1/2}(\partial\Omega) \\ u &\mapsto \gamma(u) = u|_{\partial\Omega}\,, \end{aligned}$$

and an s-dependent constant exists such that:

$$|\gamma u|_{s-1/2} \le c_s |u|_s\,. \tag{B.14}$$

(ii) For each $w \in H^{s-1/2}(\partial\Omega)$ exists an $u \in H^s(\Omega)$ with $w = \gamma u$ such that for an s-dependent constant:

$$|u|_s \le c_s |w|_{s-1/2}\,.$$

(iii) For each $w \in H^s(\Omega)$ exists a continuation $E : H^s(\Omega) \to H^s(\mathbb{R}^n)$, $w \mapsto Ew$.

Finally, an alternative definition of the H_0^k-spaces can be given:

Proposition B.21. *For $\Omega \in C^1$ and $k \in \mathbb{N}$ it is:*

$$\begin{aligned} H_0^k(\Omega) &= \{u \in H^k(\Omega) : \partial^\ell u/\partial n^\ell|_{\partial\Omega} = 0 \text{ for all } 0 \le \ell \le k-1\} \\ &= \{u \in H^k(\Omega) : D^\alpha u|_{\partial\Omega} = 0 \text{ for all } 0 \le |\alpha| \le k-1\}\,. \end{aligned}$$

B.1.4 Dual spaces

If X is a normed linear space over $\mathbb{R}$, the *dual* space X' consists of all bounded linear maps from X to $\mathbb{R}$, also written in the form: $X' = L(X, \mathbb{R})$. X' is a Banach space with the norm (*dual norm*):

$$\|x'\|_{X'} := \|x'\|_{\mathbb{R}\leftarrow X} = \sup\left\{\frac{|x'(x)|}{\|x\|_X} : 0 \ne x \in X\right\}. \tag{B.15}$$

The elements $x' \in X'$ are called *linear functionals* on X. Instead of $x'(x)$ (evaluation of x' with x) it is often written $x'(x) = \langle x, x'\rangle_{X\times X'} = \langle x', x\rangle_{X'\times X}$ in which $\langle\cdot,\cdot\rangle_{X\times X'}$ is called the *dual form* on $X \times X'$.

Lemma B.22. *Given a Banach space X being dense and continuously embedded in another Banach space Y, then the dual space Y' is continuously embedded in X'.*

Lemma B.23. *Let X and Y be normed and $T \in L(X, Y)$. For each $y' \in Y'$ the equation:*

$$\langle Tx, y'\rangle_{Y\times Y'} = \langle x, x'\rangle_{X\times X'} \quad \text{for all } x \in X$$

a unique $x' \in X'$. The linear map $y' \mapsto x'$ defines the dual operator $T' : Y' \to X'$ with $T'y' = x'$. It is $T' \in L(Y', X')$ and

$$\|T'\|_{X' \leftarrow Y'} = \|T\|_{Y \leftarrow X} .$$

If X is a Hilbert space over $\mathbb{R}$, every $y \in X$ defines by:

$$f_y(x) := (x, y)_X ,$$

a linear functional $f_y \in X'$ with $\|f_y\|_{X'} = \|y\|_X$. The converse is equally true.

Proposition B.24. (Riesz representation) *Consider a Hilbert space X and a functional $f \in X'$. Then a unique $y_f \in X$ exists such that:*

$$f(x) = (x, y)_X \quad \text{for all } x \in X \qquad \text{and } \|f\|_{X'} = \|y_f\|_X .$$

Corollary B.25. (Riesz isomorphism) *Let X be a Hilbert space.*

(i) A unique correlation $J_X \in L(X, X')$ wit $J_X y = f_y$, $J_X^{-1} f = y_f$ exists, the so-called Riesz isomorphism, which preserves the norm: $\|J_X\|_{X' \leftarrow X} = \|J_X^{-1}\|_{X \leftarrow X'} = 1$.

(ii) X' is a Hilbert space with inner product $(x', y')_{X'} = (J_X^{-1x'}, J_X^{-1} y)_X$. The dual norm $\|x'\|_{X'}$ from (B.15) agrees with the norm induces by $(x', x')_X^{1/2}$.

(iii) X can be identified via X'' via $x(x') := x'(x)$ and hereby follows:

$$J_{X'} = J_X^{-1} , J_X = J_X' \quad \text{and} \quad T'' = T \ \text{ for } T \in L(X, Y) ,$$

if also $Y = Y''$ is a Hilbert space.

(iv) One can identify X and X' by: $X = X'$, $J_X = I$.

Lemma B.26. *Consider two Hilbert spaces $V \subset U$ with a continuous and dense embedding, then U' is continuously and densely embedded in V'.*

According to Corollary B.25 (iv) U and U' can be identified, leading to the *Gelfand triple*:

$$V \subset U \subset V' \quad (V \subset U \text{ continuously and densely embedded}) .$$

and it can be shown that V and U are continuously and densely embedded in V' also. In the context of Sobolev spaces the case $U := L^2(\Omega)$ is frequently considered, yielding the following embeddings:

$$H_0^s(\Omega) \subset L^2(\Omega) \subset (H_0^s(\Omega))' , \quad s \geq 0 , \tag{B.16}$$

$$H^s(\Omega) \subset L^2(\Omega) \subset (H^s(\Omega))', \quad s \geq 0\,; \tag{B.17}$$

(B.16) and (B.17) are Gelfand triples.

Definition B.27. (Negative norm Sobolev spaces) *Let $s \geq 0$. The dual space of $H_0^s(\Omega)$ is denoted by $H^{-s}(\Omega)$ or $H_0^{-s}(\Omega)$:*

$$H_0^{-s}(\Omega) := H^{-s}(\Omega) := (H_0^s(\Omega))'\,.$$

A norm related to $H^{-s}(\Omega)$ is according to (B.15) *for all $u \in L^2(\Omega)$ defined by:*

$$\|u\|_{-s} := \sup_{0 \neq v \in H_0^s(\Omega)} \left\{ \frac{(u,v)_{L^2(\Omega)}}{\|v\|_s} \right\}, \tag{B.18}$$

in which $(u,v)_{L^2(\Omega)}$ denotes the dual form on $H_0^s(\Omega) \times H^{-s}(\Omega)$. The closure of $L^2(\Omega)$ with respect to $\|\cdot\|_{-s}$ is denoted with $H^{-s}(\Omega)$.

Finally, for spaces and norms the following relations are summarised:

$$\ldots \supset H^{-2}(\Omega) \supset H^{-1}(\Omega) \supset \quad L^2(\Omega) \quad = H^0(\Omega) \supset H^1(\Omega) \supset H^2(\Omega) \supset \ldots$$

$$\ldots \leq \quad \|u\|_{-2} \quad \leq \quad \|u\|_{-1} \quad \leq \|u\|_{L^2(\Omega)} = \quad \|u\|_0 \quad \leq \quad \|u\|_1 \quad \leq \quad \|u\|_2 \quad \leq \ldots$$

B.1.5 Functional derivatives

The *Fréchet derivative* is a derivative defined on Banach spaces. It is commonly used to generalise the derivative of a real-valued function of a single variable to the case of a vector-valued function of multiple real variables, and to define the *functional derivative* used widely in the calculus of variations.

Definition B.28. *Let V and W be Banach spaces, and $U \subset V$ be an open subset of V. A function $f : U \to V$ is called* Fréchet differentiable *at $x \in U$ if there exists a bounded linear operator $A : V \to W$ such that:*

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - Ah\|_W}{\|h\|_V} = 0\,.$$

The limit here is meant in the usual sense of a limit of a function defined on a metric space, using V and W as the two metric spaces, and the above expression as the function of argument h in V. As a consequence, it must exist for all sequences $\{h_n\}_{n=1}^{\infty}$ of non-zero elements of V which converge to the zero vector $h_n \to 0$. Equivalently, the first-order expansion holds, in Landau notation given as:

$$f(x+h) = f(x) + Ah + o(\|h\|_V)\,. \tag{B.19}$$

If there exists such an operator A, it is unique, so write $Df(x) = A$ and call it the Fréchet derivative of f at x.

B.2 Review of Agmon-Douglis-Nirenberg theory

In order to derive a priori estimates for the first integral formulation in Sec. 4.3.1 the underpinning theory by Agmon et al. [2, 3] is reviewed. Consider the general boundary value problem to find a set of scalar functions $u = \{u_1, u_2, \ldots, u_M\} \in X(\Omega)$ under conditions:

$$\mathcal{L}(x, D)u = f \quad \text{in } \Omega\,, \quad f \in Y(\Omega)\,, \tag{B.20}$$

$$\mathcal{B}(x, D)u = g \quad \text{on } \partial\Omega\,, \quad g \in B(\partial\Omega)\,, \tag{B.21}$$

involving an arbitrary differential field operator:

$$\mathcal{L}_{ij}(x, D) = \sum_{|\alpha| \le r_{ij}} a_\alpha(x) D^\alpha\,, \quad i, j = 1, \ldots, M\,, \tag{B.22}$$

and a corresponding boundary operator:

$$\mathcal{B}_{lj}(x, D) = \sum_{|\beta| \le q_{lj}} b_\beta(x) D^\beta\,, \quad l = 1, \ldots, L;\ j = 1, \ldots, M\,. \tag{B.23}$$

Here, α and β are multi-indices, r_{ij} and q_{lj} non-negative integers and D^α is a partial derivative operator, i.e. the following definitions hold for a d-dimensional problem:

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)\,, \qquad D^\alpha = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \ldots \partial_d^{\alpha_d}\,,$$
$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d\,,$$

in which the abbreviation $\partial_i^{\alpha_i} := \partial^{\alpha_i} / \partial x_i^{\alpha_i}$ is used. Note, that definitions (B.22) and (B.23) remain valid even if the differential operator D is replaced by a vector $\vec{\xi} \in \mathbb{R}^d$ and D^α by $\vec{\xi}^\alpha := \xi_1^{\alpha_1} \xi_2^{\alpha_2} \ldots \xi_d^{\alpha_d}$, in which case $\mathcal{L}(x, \vec{\xi})$ is called the *symbol* of $\mathcal{L}$.

Definition B.29. (**ADN ellipticity**) *The system* (B.20), (B.21) *is* ADN-elliptic *if there exist integer weights* $\{s_i\}$ *and* $\{t_j\}$ *for the equations and unknowns, respectively, such that:*

(i) $\deg \mathcal{L}_{ij}(x, \vec{\xi}) \le s_i + t_j$

(ii) $\mathcal{L}_{ij}(x, \vec{\xi}) \equiv 0$, *whenever* $s_i + t_j < 0$

(iii) $\det \mathcal{L}^p_{ij}(x, \vec{\xi}) \neq 0$ *for all* $\vec{\xi} \neq 0$, *where the* principle part $\mathcal{L}^p$ *of* $\mathcal{L}$ *is defined as all terms* $\mathcal{L}_{ij}$ *for which* $\deg \mathcal{L}_{ij}(x, \vec{\xi}) = s_i + t_j$.

The operator $\mathcal{L}$ is said to be uniformly elliptic *of order $2m$, if an integer $m \geq 1$ exists, with* $\deg(\det \mathcal{L}^p(x,\vec{\xi})) = 2m$*, and a positive constant C, such that:*

$$C^{-1}|\vec{\xi}|^{2m} \leq |\det \mathcal{L}^p(x,\vec{\xi})| \leq C|\vec{\xi}|^{2m}\,. \tag{B.24}$$

Consider now an ADN-elliptic system (B.20), (B.21) being uniformly elliptic of order $2m$ in line with the above definition. A minimal condition on the operators $\mathcal{L}$ and $\mathcal{B}$ to guarantee a well-posed system is that the number of rows in $\mathcal{B}$ equals m, i.e. $L = m$. An additional condition on $\mathcal{L}$ which is satisfied for all elliptic systems in three or more space conditions, but must be assumed in two dimensions, is the so-called *supplementary condition* [37].

Definition B.30. (**Supplementary Condition**) *The operator $\mathcal{L}$ satisfies the* supplementary condition *if, for every pair of linearly independent vectors $\vec{\xi}$ and $\vec{\xi}'$, the polynomial* $\det \mathcal{L}^p(x,\vec{\xi} + \tau\vec{\xi}')$ *in the complex variable τ has exactly m roots with positive imaginary part.*

An ADN-elliptic system that satisfies the Supplementary Condition is also called *regular elliptic* [37]. In addition to the sets $\{s_i\}$, $\{t_j\}$, a further set of integer weights $\{r_l\}$, $l = 1,\dots L$ is introduced for the boundary operator $\mathcal{B}$. Each r_l is attached to the l-th boundary condition in (B.21) and must satisfy the inequality:

$$\deg \mathcal{B}_{lj}(x,\vec{\xi}) \leq r_l + t_j\,,$$

on the understanding that $B_{lj} \equiv 0$ when $r_l + t_j < 0$. The principal part $\mathcal{B}^p$ of the boundary operator is defined as all terms $\mathcal{B}_{lj}$ such that $\deg \mathcal{B}_{lj}(x,D) = r_l + t_j$. The three sets of ADN-weights can always be normalised in such a way that $s_i \leq 0$, $r_l \leq 0$ and $t_j \geq 0$. However, the weights may not be unique, even with such a normalisation; i.e., there are examples of operators L for which one can define more than one principal part $\mathcal{L}^p$ that satisfies Definition B.29 [36]. An important subset of ADN-elliptic systems is the class of Petrovski systems, see [275].

Definition B.31. *A system is* elliptic in the sense of Petrovskii *if it is ADN-elliptic and $s_1 = \cdots = s_M = 0$. If, in addition, $t_1 = \cdots = t_M$ the system is called* homogeneous elliptic.

Given an ADN-elliptic operator $\mathcal{L}$, the problem (B.20), (B.21) is well-posed if and only if the boundary operator $\mathcal{B}$ "complements" the field operator $\mathcal{L}$ in a proper way. In [3] it is shown that this is equivalent to an algebraic condition, called the *Complementing*

Condition, on the principal parts $\mathcal{L}^p$ and $\mathcal{B}^p$. To state this condition, let $\vec{n}$ denote the unit normal to $\partial\Omega$ at x, $\tau_k^+(x,\vec{\xi})$ denote the m roots of $\det \mathcal{L}^p(x,\vec{\xi}+\tau\vec{n})$ with $\vec{\xi}\perp\vec{n}$, having positive imaginary part, and:

$$M^+(x,\vec{\xi},\tau)=\prod_{k=1}^{m}\left(\tau-\tau_k^+(x,\vec{\xi})\right);$$

furthermore, let $\mathcal{L}'$ denote the adjugate matrix of $\mathcal{L}^p$[1], then the condition is:

Definition B.32. (Complementing condition) *For any point $x\in\partial\Omega$ and any real non-zero unit vector $\vec{\xi}$ tangent to $\partial\Omega$ at x, regard $M^+(x,\vec{\xi},\tau)$ and the elements of the matrix $\mathcal{B}^p(x,\vec{\xi}+\tau\vec{n})\mathcal{L}'(x,\vec{\xi}+\tau\vec{n})$ given by:*

$$\sum_{j=1}^{M}\mathcal{B}_{lj}^p(x,\vec{\xi}+\tau\vec{n})\mathcal{L}'_{jk}(x,\vec{\xi}+\tau\vec{n}),\quad l=1,\dots,m;\ k=1,\dots,M \tag{B.25}$$

as polynomials in τ. The operators $\mathcal{L}$ and $\mathcal{B}$ satisfy the Complementing Condition if and only if the rows of the latter matrix are linearly independent modulo $M^+(\vec{\xi},\tau)$; i.e.,

$$\sum_{l=1}^{m}C_l\left(\sum_{j=1}^{M}\mathcal{B}_{lj}^p(x,\vec{\xi}+\tau\vec{n})\mathcal{L}'_{jk}(x,\vec{\xi}+\tau\vec{n})\right)\equiv 0 \qquad \text{mod } M^+. \tag{B.26}$$

For brevity, in what follows the boundary value problem (B.20), (B.21) is called *elliptic* if $\mathcal{L}$ is ADN-elliptic, regular and uniformly elliptic and if $\mathcal{B}$ satisfies the Complementing Condition. The main result of ADN theory, relevant to least-squares methods, is a priori estimates that give rise to energy balances required to define well-posed least-squares principles. These estimates are stated in terms of solution and data spaces that are direct products of the scalar spaces $H^k(\Omega)$ and $H^{k\pm1/2}(\partial\Omega)$, parametrized by a non-negative integer *regularity index* q:

$$X_q=\prod_{j=1}^{M}H^{q+t_j}(\Omega),\quad Y_q=\prod_{i=1}^{M}H^{q-s_i}(\Omega),\quad B_q=\prod_{l=1}^{m}H^{q-r_l-1/2}(\partial\Omega). \tag{B.27}$$

Theorem B.33. *Let $t'=\max t_j$, $q\geq r'=\max(0,\max r_l+1)$, and assume that Ω is a bounded domain of class $C^{q+t'}$. Furthermore, assume that the coefficients of $\mathcal{L}_{ij}$ are of class $C^{q-s_i}(\overline{\Omega})$ and that the coefficients of $\mathcal{B}_{lj}$ are of class $C^{q-r_l}(\partial\Omega)$. If* (B.20), (B.21) *is elliptic, $f\in Y_q$ and $g\in B_q$, then:*

(i) every solution $u\in X_{r'}$ in fact belongs to X_q;

[1]Recall that the *adjugate* or *cofactor* matrix of A is the matrix $A'=(\det A)A^{-1}$ [37]

(ii) there is a positive constant C, independent of u, f and g, such that, for every solution $u \in X_q$,

$$\sum_{j=1}^{M} \|u_j\|_{q+t_j} \leq C \left(\sum_{i=1}^{M} \|f_i\|_{q-s_i} + \sum_{l=1}^{m} \|g_l\|_{q-r_l-1/2,\partial\Omega} + \sum_{j=1}^{M} \|u_j\|_0 \right) . \tag{B.28}$$

Moreover, if the problem (B.20), (B.21) *has a unique solution in the indicated function spaces, then the $L^2(\Omega)$-norm on the right-hand side of* (B.28) *can be omitted.*

While the above result is helpful for deriving a priori estimates for elliptic differential equations, even stronger assertions can be proved concerning the properties of the corresponding differential operators; the following theorem, sometimes also called the "Equivalence Theorem" on elliptic differential equations, is cited from [207] and originally taken from [269]:

Theorem B.34. *([207], Lemma 2) Under the assumptions of the ADN-theorem B.33 the following statements are equivalent:*

(i) The problem (B.20), (B.21) *is elliptic.*

(ii) The differential operator:

$$\begin{aligned} T_q : X_q &\to Y_q \times B_q \\ u &\mapsto (\mathcal{L}u, \mathcal{B}u) \end{aligned}$$

is a Fredholm operator, i.e., the range $\operatorname{ran}(T_q) := \{T_q x \,|\, x \in X_q\} \subseteq Y_q \times B_q$ *of T_q is a finite-dimensional closed subspace of $Y_q \times B_q$ and the null-space or kernel* $\ker(T_q) := \{x \in X_q : T_q x = 0\}$ *is finite-dimensional.*

(iii) The a priori estimate (B.28) *holds for $u \in X_q$.*

In the general case the solution of problem (B.20), (B.21) is non-unique but, according to Theorem B.34, the mapping T_q is Fredholm and thereby the null space is independent of q and finite-dimensional, $\dim\ker(T_q) =: \alpha < \infty$. This allows the problem to be complemented by a number of α linear functionals Λ_j in order to obtain the following refined estimate:

Corollary B.35. *If the continuous linear functionals $\Lambda_j : X_q(\Omega) \to \mathbb{R}$, $j = 1, \ldots, \alpha$ are fixed such that they are linearly independent on* $\ker(T_q)$, *then* (B.28) *yields the stronger estimate:*

$$\sum_{j=1}^{M} \|u_j\|_{q+t_j} \leq C \left(\sum_{i=1}^{M} \|f_i\|_{q-s_i} + \sum_{l=1}^{m} \|g_l\|_{q-r_l-1/2,\partial\Omega} + \sum_{j=1}^{\alpha} |\Lambda_j u| \right) . \tag{B.29}$$

In the following, some further results related to the *adjoint* of elliptic systems are provided which will be important later on; here, the focus lies on differential operators $\mathcal{L}$ being elliptic in the sense of Petrovskii according to Definition B.31. Basically, the complementing condition implies that the m rows of matrix $\mathcal{B}$ are linearly independent for all $x \in \partial\Omega$ so that it can be completed to form a smooth, square and invertible Dirichlet matrix simply by adding further $M-m$ rows, denoted by the non-unique operator $\mathcal{C}_{ij}$, $i = 1, \ldots, M-m$; $j = 1, \ldots, M$ [207]. Under these conditions Roitberg and Seftel [207] show that a matrix operator:

$$\mathcal{B}'_{lj}(x, D) = \sum_{|\beta| \leq q'_{lj}} b_\beta(x) D^\beta \,, \quad l = 1, \ldots, M-m;\; j = 1, \ldots, M \,,$$

of the same size as $\mathcal{C}$ exists, such that for all $u, v \in X_q$ the Green's formula:

$$(\mathcal{L}u, v)_0 - (u, \mathcal{L}^* v)_0 = (\mathcal{B}u, \mathcal{C}'v)_{0,\partial\Omega} - (\mathcal{C}u, \mathcal{B}'v)_{0,\partial\Omega} \,, \tag{B.30}$$

holds, in which $\mathcal{C}'_{i,j}$, $i = 1, \ldots, m$; $j = 1, \ldots, M$ completes $\mathcal{B}'$ to form a square Dirichlet matrix and $\mathcal{L}^*$ denotes the formal adjoint of $\mathcal{L}$.

Green's formula (B.30) leads to the natural definition of the following problem, adjoint to (B.20), (B.21):

$$\mathcal{L}^*(x, D)u = f' \quad \text{in } \Omega \,, \tag{B.31}$$

$$\mathcal{B}'(x, D)u = g' \quad \text{on } \partial\Omega \,. \tag{B.32}$$

If problem (B.20), (B.21) is elliptic, the number of rows of matrix $\mathcal{B}$ is equal to m while the number of rows of $\mathcal{B}'$ is equal to $M-m$ which is generally not equal to m. Therefore problem (B.31), (B.32) may fail to be elliptic; however, in the case of Petrovskii-elliptic systems both row numbers are equal to m so that ellipticity is guaranteed as proven by Roitberg and Seftel [207]. Moreover, the solubility of (B.20), (B.21) can be expressed in terms of the adjoint system, summarised in the following theorem:

Theorem B.36. *Suppose that the operator* (B.20) *is elliptic in the sense of Petrovskii and that the boundary operator* (B.21) *can be completed to a Dirichlet system. If now the system* (B.20), (B.21) *is elliptic as a whole, then the adjoint system* (B.31), (B.32) *is elliptic as well. Furthermore,* $(f, g) \in \operatorname{ran}(T_q)$ *if and only if* $(f, g) \in Y_q \times B_q$ *such that:*

$$(f, v)_0 - (g, \mathcal{C}'v)_{0,\partial\Omega} = 0 \,,$$

for all $v \in \ker(T_q^*)$.

B.3 Auxiliary problem and regularity considerations

In order to obtain the optimal L^2-error estimates in Sec. 4.3.3(b) the following auxiliary problem is investigated in more detail: given $f \in L^2(\Omega)^4$, find $w \in V$ such that:

$$a(w,v) = (f,v)_0 \quad \text{for all } v \in V\,, \tag{B.33}$$

involving the closed subspace $V \subset H^1(\Omega)^4$ according to (4.52). The regularity of the solution w requires investigation. For this purpose consider the case $\Lambda = 0$, for which formal partial integration of (B.33) shows:

$$\begin{aligned} a(w,v) = (\mathcal{L}w, \mathcal{L}v)_0 &= -(\mathcal{L}^T\mathcal{L}w, v)_0 + (\mathcal{L}w, \mathcal{L}(x,\vec{n})v)_{0,\partial\Omega} && \text{(B.34)}\\ &= -(\mathcal{L}^T\mathcal{L}w, v)_0 + (\mathcal{C}'\mathcal{L}w, \mathcal{B}v)_{0,\partial\Omega} - (\mathcal{B}'\mathcal{L}w, \mathcal{C}v)_{0,\partial\Omega} && \text{(B.35)}\\ &= (\mathcal{L}^*\mathcal{L}w, v)_0 - (\mathcal{B}'\mathcal{L}w, \mathcal{C}v)_{0,\partial\Omega}\,. && \text{(B.36)} \end{aligned}$$

In the above, the form of (B.35) is due to the application of the Green's formula (B.30), the middle term of which is eliminated by the boundary condition $\mathcal{B}v = 0$ giving (B.36). The formal adjoint of $\mathcal{L}$ is in this case obviously given by $\mathcal{L}^* := -\mathcal{L}^T$. Thus, problem (B.33) naturally corresponds to the second-order problem:

$$\begin{cases} \mathcal{L}^*\mathcal{L}w = f\,, \\ \quad \mathcal{B}w = 0\,, \\ \mathcal{B}'\mathcal{L}w = 0\,, \end{cases} \tag{B.37}$$

in which the last condition appears in (B.33) as a natural boundary condition. This leads to a mapping of the form:

$$\begin{aligned} P_s : H^{s+2}(\Omega)^n \to H^s(\Omega)^n \times H^{s+3/2}(\partial\Omega)^m \times H^{s+1/2}(\partial\Omega)^m =: Z_s\,,\\ T_s w \mapsto (\mathcal{L}^*\mathcal{L}w, \mathcal{B}w, \mathcal{B}'\mathcal{L}w)\,, \quad s \geq 0\,. \end{aligned}$$

A straightforward calculation shows that the second-order matrix operator $\mathcal{L}^*\mathcal{L}$ is strongly elliptic. Furthermore, the boundary operators $(\mathcal{B}, \mathcal{B}'\mathcal{L})$ satisfy the complementing condition. By equivalent characterizations of the ellipticity up to the boundary, Theorem B.34, the operator $P_s : H^{s+2}(\Omega)^n \to Z_s$ is Fredholm for $s \geq 0$. These assertions are included in the following more complete result [215]:

Lemma B.37. *Let $s \geq 0$. Then the operator $P_s : H^{s+2}(\Omega)^n \to Z_s$ is a Fredholm operator with vanishing index.*

Denote the adjoint of Λ by $\tilde{\Lambda}$, then by definition $\tilde{\Lambda} : \mathbb{R}^\alpha \to H^1(\Omega)$ with:

$$\Lambda u \cdot c = (u|\tilde{\Lambda} c)_0 \,,$$

for all $u \in H^1(\Omega)^4$, $c \in \mathbb{R}^\alpha$. The mapping $\tilde{\Lambda}$ is linear and continuous and it can be stated that:

Theorem B.38. *Let $f \in L^2(\Omega)^4$ be given and let Λ be $H^1(\Omega)$-continuous. Then problem* (B.33) *has a unique solution $w \in V$. For this solution the regularity $w \in H^2(\Omega)^4$ together with the estimate:*

$$\|w\|_2 \leq c\|f\|_0 \,, \tag{B.38}$$

hold. Moreover, the following relations are true:

$$\mathcal{L}^*\mathcal{L}w + \tilde{\Lambda}\Lambda w = f \,, \tag{B.39}$$

$$\mathcal{B}w = 0 \,, \quad \mathcal{B}'\mathcal{L}w = 0 \,. \tag{B.40}$$

Proof. Obviously the variational equation (B.33) is valid for all $v \in V$ and therefore also for:

$$\phi \in \left\{ v \in [H^1(\Omega)]^4 \;\middle|\; \mathcal{B}v = 0,\, \mathcal{B}'\mathcal{L}v = 0 \text{ on } \partial\Omega \right\} .$$

Similar to (B.36) but now with $\Lambda \neq 0$, integration by parts yields:

$$(f, \phi)_0 = a(w, \phi) = (\mathcal{L}^*\mathcal{L}w, \phi)_0 + (\tilde{\Lambda}\Lambda w, \phi)_0 \,,$$

in which all boundary terms vanish, leading to (B.39). To explore the regularity property further note that (B.33) implies:

$$(\mathcal{L}w, \mathcal{L}\phi)_0 + (w, \phi)_0 = (f_w, \phi)_0 \quad \text{for all } \phi \in V \,, \tag{B.41}$$

where $f_w = f - \tilde{\Lambda}\Lambda w + w$. By Lemma B.37 the mapping:

$$\begin{aligned} \tilde{P}_s : H^{s+2}(\Omega)^4 &\to H^s(\Omega)^4 \times H^{s+3/2}(\partial\Omega)^2 \times H^{s+1/2}(\partial\Omega)^2 \,, \\ u &\mapsto (\mathcal{L}^*\mathcal{L}u + u, \mathcal{B}u, \mathcal{B}'\mathcal{L}u) \,, \quad s \geq 0 \,. \end{aligned}$$

is Fredholm with vanishing index. But this mapping has a zero null space since $u \in \ker(\tilde{P}_s)$ implies by Green's formula, (B.30), that:

$$0 = (u, \mathcal{L}^*\mathcal{L}u + u)_0 = \|\mathcal{L}u\|_0^2 + \|u\|_0^2 \,.$$

Consequently $\tilde{P}_s$ is bijective and:

$$\|u\|_{s+2} \leq c(\|\mathcal{L}^*\mathcal{L}u + u\|_s + \|\mathcal{B}u\|_{\partial\Omega,s+3/2} + \|\mathcal{B}^*\mathcal{L}u\|_{\partial\Omega,s+1/2}) \,, \tag{B.42}$$

for all $u \in H^{s+2}(\Omega)^4$, $s \geq 0$ and a constant $c = c(s)$. Thus, for a fixed f_w, there exists a unique solution $w_0 \in H^2(\Omega)^4$ such that:

$$\begin{cases} \mathcal{L}^*\mathcal{L}w_0 + w_0 = f_w \,, \\ \mathcal{B}w_0 = 0 \,, \\ \mathcal{B}'\mathcal{L}w_0 = 0 \,. \end{cases} \tag{B.43}$$

Moreover, via the homogeneous boundary conditions in (B.43) and inequality (B.42) the estimate:

$$\|w_0\|_2 \leq c\|f_w\|_0 \,, \tag{B.44}$$

is valid. From the regularity $w_0 \in H^2(\Omega)^4$ applying Green's formula leads to:

$$(\mathcal{L}w_0, \mathcal{L}\phi)_0 + (w_0, \phi)_0 = (f_w, \phi)_0 \,, \tag{B.45}$$

for all $\phi \in V$. Writing $\phi = w - w_0$, equations (B.41) and (B.45) imply:

$$\|\mathcal{L}(w - w_0)\|_0^2 + \|w - w_0\|_0^2 = 0 \,, \tag{B.46}$$

and thus $w = w_0$. From (4.36) and (B.33) it follows:

$$\|w\|_1 \leq c\|f\|_0 \,, \tag{B.47}$$

which, by the definition of f_w, yields:

$$\begin{aligned} \|f_w\|_0 = \|f - \tilde{\Lambda}\Lambda w + w\|_0 &\leq \|f\|_0 + \|\tilde{\Lambda}\Lambda w\|_0 + \|w\|_0 \\ &\leq \|f\|_0 + c\|w\|_1 + \|w\|_0 \leq \tilde{c}\|f\|_0 \,. \end{aligned} \tag{B.48}$$

The first inequality in (B.48) follows due to the continuity of Λ and $\tilde{\Lambda}$, such that:

$$\|\tilde{\Lambda}\Lambda w\|_0 \leq \|\tilde{\Lambda}\Lambda w\|_1 \lesssim \|\Lambda w\|_E \lesssim \|w\|_1$$

and the second inequality from $\|w\|_0 \leq \|w\|_1$ and (B.47). Now the assertions (B.38) - (B.40) follow via (B.43), (B.44) and (B.48). □

Consider next the Galerkin approximation of the solution $w \in V \cap H^2(\Omega)$ to problem (B.33), which is denoted by $w_h \in V_h$ and given through:

$$a(w_h, v) = (f, v)_0 \,, \quad \text{for all } v \in V_h \,. \tag{B.49}$$

The existence of a unique solution is guaranteed for small parameters h for the same reason as for the least squares approximation. To derive the error estimates, firstly observe that via (B.30) and (B.39), (B.40) the following identities:

$$\begin{aligned} a(w,v) &= (\mathcal{L}w, \mathcal{L}v)_0 + \Lambda w \Lambda v \\ &= (\mathcal{L}^*\mathcal{L}w, v)_0 + (\tilde{\Lambda}\Lambda w, v)_0 + (\mathcal{C}'\mathcal{L}w, \mathcal{B}v)_{0,\partial\Omega} - (\mathcal{B}'\mathcal{L}w, \mathcal{C}v)_{0,\partial\Omega} \\ &= (f,v)_0 + (C'\mathcal{L}w, \mathcal{B}v)_{0,\partial\Omega} \end{aligned} \tag{B.50}$$

are valid for all $v \in H^1(\Omega)^4$. Accordingly, (B.49) and (B.50) imply:

$$a(w - w_h, v) = (\mathcal{C}'\mathcal{L}w, \mathcal{B}v)_{0,\partial\Omega} \quad \text{for all } v \in V_h\,. \tag{B.51}$$

On the other hand, for all $\psi \in V_h$:

$$\begin{aligned} a(\psi - w_h, \psi - w_h) &= a(\psi - w, \psi - w_h) + a(w - w_h, \psi - w_h) \\ &\leq c\|\psi - w\|_1 \|\psi - w_h\|_1 + a(w - w_h, \psi - w_h)\,. \end{aligned} \tag{B.52}$$

By the coerciveness (4.54) for small h, the following inequality is obtained using (B.52):

$$\|\psi - w_h\|_1 \leq c\left(\|\psi - w\|_1 + \sup_{0\neq\gamma\in V_h} \frac{a(w - w_h, \gamma)}{\|\gamma\|_1}\right),$$

yielding:

$$\|w - w_h\|_1 \leq c\left(\inf_{\psi\in V_h} \|\psi - w\|_1 + \sup_{0\neq\gamma\in V_h} \frac{a(w - w_h, \gamma)}{\|\gamma\|_1}\right). \tag{B.53}$$

The above inequality essentially corresponds to the Lemma of Berger, Scott and Strang, see [44]. Obviously the first term on the right-hand side is bounded by the error of the best approximation in the associated finite element space, that is Theorem 4.11 gives:

$$\inf_{\psi\in V_h} \|\psi - w\|_1 \leq ch\|w\|_2\,.$$

Via (B.51) the last term in (B.53) can be simplified to:

$$\begin{aligned} a(w - w_h, \gamma) &= (\mathcal{C}'\mathcal{L}w, \mathcal{B}\gamma)_{0,\partial\Omega} \\ &\leq c\|w\|_{1,\partial\Omega}\|\gamma\|_{0,\partial\Omega} \leq \tilde{c}\|w\|_2 h^{3/2}\|\gamma\|_1\,, \end{aligned}$$

in which the $H^1(\partial\Omega)$-norm is estimated via the trace inequality (B.14) and the $L^2(\partial\Omega)$-norm via assumption (4.44). Thus (B.53) implies convergence via:

$$\|w - w_h\|_1 \leq ch\|w\|_2\,. \tag{B.54}$$

B.4 Auxiliary problem in an infinite sector

Below the first integral boundary value problem (4.75), (4.76) in a sector domain Q is considered in more detail. The corner singularities and related β-values are investigated for various opening angles θ and, for completeness, also shift theorems are provided in the weighted Sobolev spaces defined in Sec. 4.3.4 finally leading to Theorem 4.16.

It is instructive to start with the geometry of an infinite sector, i.e. (4.72)-(4.74) with $R \to \infty$. The non-dimensioned first integral of the Stokes equations in terms of the streamfunction Ψ and the potential field Φ can, according to (1.45), be written as:

$$\frac{\partial^2 \Psi}{\partial y^2} - \frac{\partial^2 \Psi}{\partial x^2} + 2\frac{\partial^2 \Phi}{\partial x \partial y} = f_1 \,, \tag{B.55a}$$

$$\frac{\partial^2 \Phi}{\partial y^2} - \frac{\partial^2 \Phi}{\partial x^2} - 2\frac{\partial^2 \Psi}{\partial x \partial y} = f_2 \,, \tag{B.55b}$$

which can formally be identified as a plane elasticity system:

$$\Delta v_1 - 2\frac{\partial}{\partial x}\left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y}\right) = f_1 \,, \tag{B.56a}$$

$$\Delta v_2 - 2\frac{\partial}{\partial y}\left(\frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y}\right) = f_2 \,, \tag{B.56b}$$

with displacement components $v_1 = \Psi$, $v_2 = -\Phi$ and Lamé coefficients $\mu = -1$, $\lambda = 3$. This identification seems crude from a physical point of view due to the signs of the Lamé coefficients and particularly the, at least in this context, anomalous boundary conditions stemming from the fluid mechanics context. These are given by $\mathcal{B}u = [g^0, g^1]$ with:

$$\mathcal{B}v\big|_{\Gamma_0} = \begin{pmatrix} 1 & 0 \\ \partial_n & 0 \end{pmatrix} v = g^0 \quad \text{and} \quad \mathcal{B}v\big|_{\Gamma_1} = \begin{pmatrix} 0 & 1 \\ 0 & \partial_n \end{pmatrix} v = g^1 \,, \tag{B.57}$$

in which the complete sector boundary is given by $\partial Q = S_0 \cup S_1 = \Gamma_0 \cup \Gamma_1$ and three cases are distinguished: (i) $\Gamma_0 = S_0 \cup S_1$, $\Gamma_1 = \emptyset$, (ii) $\Gamma_0 = \emptyset$, $\Gamma_1 = S_0 \cup S_1$, and (iii) $\Gamma_0 = S_0$, $\Gamma_1 = S_1$; the completed boundary value problem is abbreviated by:

$$\mathcal{L}v = f \quad \text{in } Q \,, \tag{B.58}$$

$$\mathcal{B}v = [g^0, g^1] \quad \text{on } \partial Q \,. \tag{B.59}$$

Note that the variational principle established in [230] is a direct consequence of well-known results in elasticity theory [114]. However, the above analogy is more mathematically than physically useful in that a considerable part of the analysis for the classical elasticity system can be extended to the present boundary value problem. In the following, results from Guo and Babuska [114] are generalised to obtain estimates in the infinite sector Q.

Let $\bar{v} = (v_r, v_\varphi) = Av$, $\bar{f} = Af$ with:

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix},$$

then equations (B.56a), (B.56b) in polar coordinates are:

$$\Delta v_r - \frac{1}{r^2} v_r - \frac{2}{r^2}\frac{\partial v_\varphi}{\partial \varphi} - 2\frac{\partial}{\partial r}\left(\frac{\partial v_r}{\partial r} + \frac{v_r}{r} + \frac{1}{r}\frac{\partial v_\varphi}{\partial \varphi}\right) = f_r\,, \tag{B.60}$$

$$\Delta v_\varphi - \frac{1}{r^2} v_\varphi + \frac{2}{r^2}\frac{\partial v_r}{\partial \varphi} - \frac{2}{r}\frac{\partial}{\partial \varphi}\left(\frac{\partial v_r}{\partial r} + \frac{v_r}{r} + \frac{1}{r}\frac{\partial v_\varphi}{\partial \varphi}\right) = f_\varphi\,, \tag{B.61}$$

and additionally one of three kinds of boundary conditions, may be imposed on ∂Q:

$$\bar{\mathcal{B}}_0 \bar{v}\big|_{\varphi=0,\theta} = \begin{pmatrix} \cos(\varphi)\partial_r & -\sin(\varphi)\partial_r \\ -\sin(\varphi) + \cos(\varphi)\partial_\varphi & -\cos(\varphi) - \sin(\varphi)\partial_\varphi \end{pmatrix} \bar{v} = \bar{g}^0 = g^0\,, \tag{B.62}$$

$$\bar{\mathcal{B}}_1 \bar{v}\big|_{\varphi=0,\theta} = \begin{pmatrix} \sin(\varphi)\partial_r & \cos(\varphi)\partial_r \\ \cos(\varphi) + \sin(\varphi)\partial_\varphi & -\sin(\varphi) + \cos(\varphi)\partial_\varphi \end{pmatrix} \bar{v} = \bar{g}^1 = g^1\,, \tag{B.63}$$

$$\bar{\mathcal{B}}_0 \bar{v}\big|_{\varphi=0} = \bar{g}^0 \quad \text{and} \quad \bar{\mathcal{B}}_1 \bar{v}\big|_{\varphi=\omega} = \bar{g}^1\,, \tag{B.64}$$

in which (B.62) corresponds to velocity, (B.63) to stress and (B.64) to mixed boundary conditions. The boundary value problem is written as: $[\bar{\mathcal{L}}, \bar{\mathcal{B}}]\bar{v} = [\bar{f}, \bar{g}^0, \bar{g}^1]$. By change of variable $\tau = \ln(1/r)$ the sector is converted to a strip domain $D = \{(\tau, \varphi) \,|\, -\infty < \tau < \infty,\, 0 < \varphi < \theta\}$ and the system $\tilde{L}\tilde{v} = \tilde{f}$ is obtained:

$$\tilde{v}_\tau - \frac{\partial^2 \tilde{v}_\tau}{\partial \tau^2} + \frac{\partial^2 \tilde{v}_\tau}{\partial \varphi^2} + 2\frac{\partial^2 \tilde{v}_\varphi}{\partial \tau \partial \varphi} = \tilde{f}_\tau\,, \tag{B.65}$$

$$-\tilde{v}_\varphi + \frac{\partial^2 \tilde{v}_\varphi}{\partial \tau^2} - \frac{\partial^2 \tilde{v}_\varphi}{\partial \tau^2} - 2\frac{\partial^2 \tilde{v}_\tau}{\partial \tau \partial \varphi} = \tilde{f}_\varphi\,, \tag{B.66}$$

with boundary conditions of the form $\tilde{B}_l \tilde{v} = \tilde{g}^l$, where:

$$\tilde{\mathcal{B}}_0 \tilde{v}\big|_{\varphi=0,\theta} = \begin{pmatrix} -\cos(\varphi)\partial_\tau & \sin(\varphi)\partial_\tau \\ -\sin(\varphi) + \cos(\varphi)\partial_\varphi & -\cos(\varphi) - \sin(\varphi)\partial_\varphi \end{pmatrix} \tilde{v} = \tilde{g}^0 = e^{-\tau} g^0\,, \tag{B.67}$$

$$\tilde{\mathcal{B}}_1 \tilde{v}\big|_{\varphi=0,\theta} = \begin{pmatrix} -\sin(\varphi)\partial_\tau & -\cos(\varphi)\partial_\tau \\ \cos(\varphi) + \sin(\varphi)\partial_\varphi & -\sin(\varphi) + \cos(\varphi)\partial_\varphi \end{pmatrix} \tilde{v} = \tilde{g}^1 = e^{-\tau} g^1\,, \tag{B.68}$$

$$\tilde{v} = (\tilde{v}_\tau, \tilde{v}_\varphi) := (v_r(e^{-\tau}, \varphi), v_\varphi(e^{-\tau}, \varphi))\,,$$

$$\tilde{f} = (\tilde{f}_\tau, \tilde{f}_\varphi) := e^{-2\tau}(v_r(e^{-\tau}, \varphi), v_\varphi(e^{-\tau}, \varphi))\,,$$
$$\tilde{g} = (\tilde{g}_\tau, \tilde{g}_\varphi) := e^{-\tau}(g^0(e^{-\tau}, \varphi), g^1(e^{-\tau}, \varphi))\,.$$

Finally, by Fourier transform $\hat{v} = (\hat{v}_\tau, \hat{v}_\varphi) = (F(\tilde{v}_\tau), F(\tilde{v}_\varphi))$ with:

$$F(\tilde{v}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\mathrm{i}\eta\tau} \tilde{v}(\tau, \varphi)\, \mathrm{d}\tau\,, \quad \eta = \xi + \mathrm{i}h\,,\ -\infty < \xi < \infty,\ h > 0\,,$$

a system of ODEs in the strip $I = (0, \theta)$ is obtained:

$$(1+\eta^2)\hat{v}_\tau + \frac{\mathrm{d}^2\hat{v}_\tau}{\mathrm{d}\varphi^2} + 2\mathrm{i}\eta\frac{\mathrm{d}\hat{v}_\varphi}{\mathrm{d}\varphi} = \hat{f}_\tau\,, \tag{B.69}$$

$$-(1+\eta^2)\hat{v}_\varphi - \frac{\mathrm{d}^2\hat{v}_\varphi}{\mathrm{d}\varphi^2} + 2\mathrm{i}\eta\frac{\mathrm{d}\hat{v}_\tau}{\mathrm{d}\varphi} = \hat{f}_\varphi\,, \tag{B.70}$$

with boundary conditions $\hat{\mathcal{B}}_l\hat{v} = \hat{g}^l$ on $\partial I = \{0, \theta\}$:

$$\hat{\mathcal{B}}_0\hat{v}\big|_{\varphi=0,\theta} = \begin{pmatrix} -\mathrm{i}\eta\cos(\varphi) & \mathrm{i}\eta\sin(\varphi) \\ -\sin(\varphi) + \cos(\varphi)\partial_\varphi & -\cos(\varphi) - \sin(\varphi)\partial_\varphi \end{pmatrix} \hat{v} = \hat{g}^0\,, \tag{B.71}$$

$$\hat{\mathcal{B}}_1\hat{v}\big|_{\varphi=0,\theta} = \begin{pmatrix} -\mathrm{i}\eta\sin(\varphi) & -\mathrm{i}\eta\cos(\varphi) \\ \cos(\varphi) + \sin(\varphi)\partial_\varphi & -\sin(\varphi) + \cos(\varphi)\partial_\varphi \end{pmatrix} \hat{v} = \hat{g}^1\,, \tag{B.72}$$

$$\hat{\mathcal{B}}_0\hat{v}\big|_{\varphi=0} = \hat{g}^0 \quad \text{and} \quad \hat{\mathcal{B}}_1\hat{v}\big|_{\varphi=\omega} = \hat{g}^1\,, \tag{B.73}$$

that can be written as:

$$\mathcal{U}(\eta)\hat{v} = [\hat{\mathcal{L}}(\eta), \hat{\mathcal{B}}(\eta)]\hat{v} = [\hat{f}, \hat{g}^0, \hat{g}^1]\,. \tag{B.74}$$

The operator $\mathcal{U}(\eta)$ depends polynomially on the complex parameter η. By arguments used in [114] for all η, with the exception of certain isolated points, $\mathcal{U}(\eta)$ realises an isomorphism $H^2(I) \to L^2(I) \times \mathbb{C}^2 \times \mathbb{C}^2$. Consequently, the inverse operator $\mathcal{R}(\eta) = \mathcal{U}(\eta)^{-1}$ is an operator-valued meromorphic function of η with poles of finite multiplicity; these poles are the eigenvalues of $\mathcal{U}(\eta)$ [114, 115, 149]. For each pole η of $\mathcal{R}(\eta)$, the homogeneous problem of (B.74) has at least one non-trivial solution, an eigenvector, corresponding in $H^2(\Omega)$. The transcendental equations which the eigenvalues satisfy have been derived for the elasticity and Stokes equations for instance in [114, 115], where typically the corresponding biharmonic problem is considered [34]. Since the coefficients of the operators $\hat{\mathcal{L}}$ and $\hat{\mathcal{B}}$ are constants, transcendental equations can be derived directly from the homogeneous equations (B.69)-(B.70) and the boundary conditions (B.71)-(B.73).

Lemma B.39. *Let $\eta = \mathrm{i}z$ be an eigenvalue of $\mathcal{U}(\eta)$, then:*

1. *for field equations* (B.69), (B.70) *with either pure velocity boundary conditions* (B.71) *or pure stress boundary conditions* (B.72), *z satisfies:*

$$\sin^2((z-1)\theta) = (z-1)^2 \sin^2(\theta)\,; \tag{B.75}$$

2. *for the field equations* (B.69), (B.70) *with mixed boundary conditions* (B.73), *z satisfies:*

$$\cos^2((z-1)\theta) = (z-1)^2 \sin^2(\theta)\,. \tag{B.76}$$

Proof. Since the coefficients of the differential operator $\hat{\mathcal{L}}$ and boundary operator $\hat{\mathcal{B}}$ are constants, the solution can be written as $\hat{v}_i = c_j e^{b_j \varphi} v_{i,j}$ with complex-valued constants c_j, b_j and complex eigenvectors v, in which b and c satisfy $A_{ij}(b)v_j = 0$ with:

$$A(b) = \begin{pmatrix} 1+\eta^2+b^2 & 2\mathrm{i}\eta b \\ 2\mathrm{i}\eta b & -(1+\eta^2)-b^2 \end{pmatrix}.$$

Then:

$$\begin{aligned} \det A(b) &= -[1+\eta^2+b^2]^2 + 4\eta^2 b^2 \\ &= -[(\eta-b)^2-1][(\eta+b)^2+1]\,, \end{aligned}$$

and with $\eta = iz$ it follows from $\det A(b) = 0$, that:

$$b_{1/2} = \pm\mathrm{i}(z+1)\,, \quad b_{3/4} = \pm\mathrm{i}(z-1)\,,$$

with corresponding eigenvectors:

$$v_{1/3} = \begin{pmatrix} -\mathrm{i} \\ 1 \end{pmatrix}, \quad v_{2/4} = \begin{pmatrix} \mathrm{i} \\ 1 \end{pmatrix}.$$

Considering that $\hat{v}$ must be real-valued, the solution of the homogeneous problems is:

$$\hat{v} = c_1 \begin{pmatrix} \sin(z+1)\varphi \\ \cos(z+1)\varphi \end{pmatrix} + c_2 \begin{pmatrix} \cos(z+1)\varphi \\ -\sin(z+1)\varphi \end{pmatrix} + c_3 \begin{pmatrix} \sin(z-1)\varphi \\ \cos(z-1)\varphi \end{pmatrix} + c_2 \begin{pmatrix} \cos(z-1)\varphi \\ -\sin(z-1)\varphi \end{pmatrix},$$

involving the four real-valued constants c_j. Now, inserting the above solution representa-

tion into boundary condition (B.71) results in a system $Q_{ij}(z)c_j = g$ with:

$$Q(z) = \begin{pmatrix} 0 & z & 0 & z \\ z & 0 & z-2 & 0 \\ z\sin(z\theta) & z\cos(z\theta) & z\sin((z-2)\theta) & z\cos((z-2)\theta) \\ z\cos(z\theta) & -z\sin(z\theta) & (z-2)\cos((z-2)\theta) & -(z-2)\sin((z-2)\theta) \end{pmatrix}.$$

For the existence of non-trivial solutions it is necessary and sufficient that $\det Q(z) = 0$, which leads, following some trigonometric manipulations, to:

$$\begin{aligned} \det Q(z) &= -z^2\left[2z(2-z) - 2\cos(2(z-1)\theta) + 2(z-1)^2\cos(2\theta)\right], \\ &= -4z^2[\sin^2((z-1)\theta) - (z-1)^2\sin^2(\theta)] \stackrel{!}{=} 0, \end{aligned} \tag{B.77}$$

providing the transcendental equation (B.75). The identical result is obtained for boundary conditions (B.72) while insertion of the mixed conditions (B.73) into the solution representation delivers the transcendental equation (B.76). □

From equations (B.75), (B.76) it is easy to see that zeroes of these equations are symmetric with respect to the origin and the axis $\operatorname{Re} z = 1$ in the complex plane. Hence the eigenvalues of $\mathcal{U}(\eta)$ are located in the complex plane symmetrically with respect to the origin and the axis $\operatorname{Im}\eta = 1$. Let T_η denote the eigenvalues and κ, κ_1 be positive numbers such that:

$$1 + \kappa = \kappa_1 = \min_{\substack{\eta\in T_\eta \\ \operatorname{Im}(\eta)>1}} \operatorname{Im}(\eta) = \min_{\substack{\eta\in T_\eta \\ \operatorname{Im}(\eta)\neq 1}} |\operatorname{Im}(\eta)|.$$

Next a proof of the Agranovich and Vishik conditions [4] is given. Let $D\hat{v} = \mathrm{i}(\mathrm{d}\hat{v}/\mathrm{d}\varphi)$ and $\hat{\mathcal{L}}^p(D,\eta)$ be the principal part of the operator $\hat{\mathcal{L}}(D_\varphi,\eta)$ in matrix form reading:

$$\hat{\mathcal{L}}^p(D,\eta) = \begin{pmatrix} \eta^2 - D^2 & 2\eta D \\ 2\eta D & -\eta^2 + D^2 \end{pmatrix}.$$

Lemma B.40. (Condition I) *For $\xi \in \mathbb{R}$, $\eta \in \Sigma_{\phi_1} = \{\eta : |\arg\eta| < \phi_1$ or $|\arg\eta - \pi| < \phi_1\}$ with any $\phi_1 \in (0, \pi/2)$ and $|\eta| + |\xi| \neq 0$, $\det(\hat{\mathcal{L}}^p(\xi,\eta)) \neq 0$. Furthermore, the equation $\det(\hat{\mathcal{L}}^p(\zeta,\eta)) = 0$ in ζ has equal numbers of roots in the upper and lower half-planes for $\eta \in \Sigma_{\phi_1}$ and $\eta \neq 0$.*

Proof. It is easy to see that:

$$\det\hat{\mathcal{L}}^p(\xi,\eta) = -(\xi^2+\eta^2)^2 \neq 0,$$

for $\xi \in \mathbb{R}$, $\eta \in \Sigma_{\phi_1}$ with any $\phi_1 \in (0, \pi/2)$, and $|\xi| + |\eta| \neq 0$. Also it is seen that $\zeta = \pm i\eta$ are the roots of the equation $\det(\hat{\mathcal{L}}^p(\zeta, \eta)) = 0$ in ζ (complex). Hence the equation has two roots in the upper and lower planes, respectively if $0 \neq \eta \in \Sigma_{\phi_1}$. □

Let $\hat{\mathcal{B}}^p(D, \eta)$ be the principal part of the boundary operator $\hat{\mathcal{B}}(D_\varphi, \eta)$ given by (B.71)-(B.73), and define for $\varphi = \varphi_0$ with $\varphi_0 = 0$ or θ, respectively, the two infinite strips I_{φ_0} with $I_0 = (0, \infty)$ and $I_\theta = (-\infty, \theta)$. This leads to the following lemma:

Lemma B.41. (Condition II) *For any $\phi_1 \in (0, \pi/2)$, if $\eta \neq 0$ and $\eta \in \Sigma_{\phi_1}$, the equation on the half-line:*

$$\hat{\mathcal{L}}^p(D, \eta)\hat{w} = 0\,, \quad \varphi \in I_{\varphi_0}\,, \tag{B.78}$$

$$\hat{\mathcal{B}}^p(D, \eta)\hat{w}\big|_{\varphi=\varphi_0} = \hat{h}\,, \tag{B.79}$$

has a unique stable solution $\hat{w}$ such that $|\hat{w}| \to 0$ as $\varphi \to \infty$ (resp. $\varphi \to -\infty$).

Proof. The proof of the above lemma is provided for $\varphi_0 = 0$ and $I_{\varphi_0} = (0, \infty)$ only; the proof for $\varphi_0 = \theta$ being very similar. For the homogeneous equation (B.78) the solution $\hat{w}$ must have the form $e^{b\varphi}(c_1, c_2)^T$ with b satisfying the equation:

$$\det \hat{\mathcal{L}}^p(ib, \eta) = -(b^2 - \eta^2)^2 = 0.$$

Hence $b = \pm\eta$ are the roots with multiplicity of two; for $\eta \in \Sigma_{\phi_1}$ and $\eta \neq 0$, $\operatorname{Re} b = \operatorname{Re} \eta \neq 0$. Letting $\alpha = -\operatorname{sgn}(\operatorname{Re} \eta)$:

$$\hat{w} = \begin{pmatrix} \hat{w}_\tau \\ \hat{w}_\varphi \end{pmatrix} = c_1 e^{\alpha\eta\varphi} \begin{pmatrix} 1 \\ \alpha\mathrm{i} \end{pmatrix} + c_2 e^{\alpha\eta\varphi} \begin{pmatrix} \varphi \\ \alpha\mathrm{i}\varphi \end{pmatrix},$$

is a stable solution if c_1, c_2 can be uniquely determined by any given boundary condition $\hat{h}$. Velocity boundary conditions (B.79), together with (B.71) [114] lead to:

$$\hat{\mathcal{B}}^p(D, \eta)\hat{w}\big|_{\varphi=0} = \begin{pmatrix} -\mathrm{i}\eta\hat{w}_\tau \\ \partial_\varphi \hat{w}_\tau \end{pmatrix}_{\varphi=0} = \begin{pmatrix} -\mathrm{i}\eta & 0 \\ \eta\mathrm{i} & \alpha\mathrm{i} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \hat{h}\,,$$

which is uniquely solvable for $\eta \neq 0$. The same is true for the stress boundary operator, i.e. the homogeneous system (B.78), (B.79) has a unique stable solution for $\eta \in \Sigma_{\phi_1}$ and $\eta \neq 0$, namely:

$$|\hat{w}| \leq c_0 e^{-b_0\varphi} \to 0 \quad \text{as } \varphi \to \infty\,,$$

with $c_0 > 0$ and $b_0 = -|\operatorname{Re} \eta|$. □

For further analysis, η-dependent norms in the strip domain I are required:

$$|||\hat{v}|||^2_{H^k(I)} := \sum_{l=0}^{k} |\eta|^{2l} \|\hat{v}\|^2_{H^{k-l}(I)} \,.$$

Verification of Conditions I and II above, Agranovich and Vishik [4], leads directly to the following theorem for $k = 0$, which can be generalised along the lines of Guo and Schwab [115] for the case of $k > 0$:

Theorem B.42. *Suppose there is no pole of $\mathcal{R}(\eta)$ on the line $\ell_h = \{\eta \in \mathbb{C} \,|\, \operatorname{Im}\eta = h\}$, then problem* (B.69), (B.70) *with one of the boundary conditions* (B.71)-(B.73) *and data $\hat{f} \in H^k(I)$, $\hat{g}^l \in \mathbb{C}^2$ admits a unique solution $\hat{v} = \mathcal{R}(\eta)[\hat{f}, \hat{g}^0, \hat{g}^1] \in H^{k+2}(I)$ and for $k \geq 0$ and all $\eta \in \ell_h$ the following estimate holds:*

$$|||\hat{v}|||^2_{H^{k+2}(I)} \leq c \left(|||\hat{f}|||^2_{H^k(I)} + \sum_{l,m=0,1} |\eta|^{2(k-m)+3} |\hat{g}^l_m|^2 \right) , \tag{B.80}$$

with a constant c independent of ξ.

Proof. Due to Lemmas B.40 and B.41 the Conditions I and II of Agranovich and Vishik [4] are satisfied with $\eta \in \Sigma_{\phi_1}$, $\phi_1 \in (0, \pi/2)$. Via Theorem 6.1 of [4] the above inequality holds with a constant c independent of η and $\hat{v}$ if $\eta \in \Sigma_{\phi_1}$ and $|\eta| > \eta_0$, where η_0 is some positive real constant. The line $\operatorname{Im}\eta = h$ is contained in Σ_{ϕ_1} except for a finite segment for which $|\operatorname{Re}\eta| < |h|/\sin(\phi_1)$. Hence, the proposed inequality holds for η on the line $\operatorname{Im}\eta = h$ with $|\eta| > \eta_0$. For values of η on the line with $|\eta| \leq \eta_0$, $\mathcal{R}(\eta)$ is analytic. □

Completely analogous to Lemma 4.4 of Guo and Babuska [114] the following result is obtained:

Lemma B.43. *Let $\mathcal{T}$ be a strip $\{\eta \,|\, 1 - h < \operatorname{Im}\eta < 1 + h\}$, $0 < h < \kappa$, then the resolvent $\mathcal{R}(\eta)$ has no pole in $\mathcal{T}$ for the velocity, stress and mixed problems, and the point $\eta = i$ is the only pole of $\mathcal{R}(\eta)$ in $\mathcal{T}$ for the mixed problem.*

In the following the goal is to transfer the result from Theorem B.42 back to the Cartesian coordinate system; this requires the use of technical results from [115], together with the definition of a norm on the strip domain D:

$$\|\tilde{v}\|_{H^k_h(D)^2} := \sum_{\alpha \leq k} \|e^{h\tau} D^\alpha \tilde{v}\|_{L^2(D)^2} \,. \tag{B.81}$$

Lemma B.44. *If $\bar{v}(r, \varphi) \in H^k_\beta(Q)$, $k \geq 0$, then $\tilde{v}(\tau, \varphi) := \bar{v}(e^{-\tau}, \varphi) \in H^k_h(D)$ with $h = k - 1 - \beta$, and:*

$$c_1 \|\tilde{v}\|_{H^k_h(D)} \leq \|\bar{v}\|_{H^k_\beta(Q)} \leq c_2 \|\tilde{v}\|_{H^k_h(D)} \,. \tag{B.82}$$

Moreover, for $0 \leq l \leq 1$, $\tilde{v}_l(\tau,\varphi) = e^{(l-2)\tau}\bar{v}(e^{-\tau},\varphi) \in H^k_h(D)$, *with* $h = k+1-l-\beta$, *and:*

$$c_1\|\tilde{v}_l\|_{H^k_h(D)} \leq \|\bar{v}\|_{H^k_\beta(Q)} \leq c_2\|\tilde{v}_l\|_{H^k_h(D)}\,, \tag{B.83}$$

in which the two constants c_1 *and* c_2 *are independent of* v, $\tilde{v}$. *Now, define* $D = \mathbb{R}\times I = (-\infty,\infty)\times(0,\theta)$ *and let* $\tilde{v} \in H^k_h(D)$, $k \geq 0$, *then* $\hat{v}(\eta,\cdot) = F(\tilde{v}) \in H^k(I)$, *and:*

$$c_1\|\tilde{v}\|_{H^k_h(D)} \leq \int_{-\infty+\mathrm{i}h}^{\infty+\mathrm{i}h} \|\hat{v}\|_{H^k(I)}\,\mathrm{d}\eta \leq c_2\|\tilde{v}\|_{H^k_h(D)}\,, \tag{B.84}$$

where the positive constants c_1, c_2 *are independent of* $\tilde{v}$.

Lemma B.45. *Let* $\bar{G}^l_m(r,\varphi) \in H^{k-m+2}_\beta(Q)$ *with* $\bar{G}^l|_{\Gamma^l} = \bar{g}^l$, $l = 0,1$, *and let* $\hat{G}^l = F(\tilde{G}^l)$, *with* $\tilde{G}^l = e^{-l\tau}\bar{G}(e^{-\tau},\varphi)$, *leading to the a priori estimate:*

$$|\eta|^{2(k-m)+3)}|\hat{g}^l_m|^2 \leq c\|\hat{G}^l_m\|_{H^{k-m+2}(I)}\,, \quad m,l = 0,1;\ k \geq 0\,. \tag{B.85}$$

Theorem B.46. *Let* $\bar{f} \in H^k_\beta(Q)^2$, $\bar{g}^l_m \in H^{k-m+\frac{3}{2}}_\beta(\Gamma^l)$, $m,l = 0,1$ *with* $k \geq 0$. *If* $\mathcal{R}(\eta)$ *has no pole on the line* $\ell_h = \{\eta \in \mathbb{C} \,|\, \operatorname{Im}\eta = h\}$ *with* $h = k+1-\beta$, *i.e.* $k-\kappa < \beta < k+\kappa$, *then the problem* (B.60), (B.61) *with one of the boundary conditions* (B.62)-(B.64) *admits a unique solution* $\bar{v} \in H^{k+2}_\beta(Q)$ *and:*

$$\|\bar{v}\|_{H^{k+2}_\beta(Q)} \leq C\left(\|\bar{f}\|_{H^k_\beta(Q)} + \sum_{m,l=0,1}\|\bar{g}^l_m\|_{H^{k-m+\frac{3}{2}}_\beta(\Gamma^l)}\right). \tag{B.86}$$

Proof. Via the definition $H^{k-m+3/2}_\beta(\Gamma^l)$, there exists $\bar{G}^l \in H^{k+2}_\beta(Q)\times H^{k+1}_\beta(Q)$, $l = 0,1$, such that $\bar{G}^l|_{\Gamma^l} = \bar{g}^l$, and for $l = 0,1$:

$$\frac{1}{2}\|\bar{G}^l_m\|_{H^{k-m+2}_\beta(Q)} \leq \|\bar{g}^l_m\|_{H^{k-m+3/2}_\beta(Q)} \leq \|\bar{G}^l_m\|_{H^{k-m+2}_\beta(Q)}\,. \tag{B.87}$$

Due to Lemma B.44 the partial Fourier transforms $\hat{f} \in H^k(I)$, $\hat{G}^l_m \in H^{k-m+2}(I)^2$ exist and (B.82)-(B.84) hold. Via Theorem B.42 the system (B.69), (B.70) with one of the boundary conditions (B.71)-(B.73) has a unique solution $\hat{v} \in H^{k+2}(I)^2$ for $k \geq 0$ and the a priori estimate (B.80) holds. The last term can be replaced according to Lemma B.45. Since $\mathcal{R}(\eta)$ has no pole on the line $\{\eta \,|\, \operatorname{Im}\eta = h\}$ with $h = k+1-\beta$, the solution:

$$\tilde{v} := F^{-1}(\hat{v}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty+\mathrm{i}h}^{\infty+\mathrm{i}h} e^{\mathrm{i}\eta\tau}(\hat{v})\,\mathrm{d}\eta\,,$$

of the system (B.65)-(B.66) is in $H_h^{k+2}(D)$ and by Lemma B.44 the following holds:

$$\|\tilde{v}\|_{H_h^{k+2}(D)} \leq C \left(\|\tilde{f}\|_{H_h^k(D)} + \sum_{m,l=0,1} \|\tilde{G}_m^l\|_{H_h^{k-l}(D)} \right) .$$

Consequently, $\bar{v} \in H_\beta^{k+2}(Q)^2$ is the unique solution of problem (B.60)-(B.61) with one of the boundary conditions (B.62)-(B.64) and it follows from (B.82), (B.83) that for any $k \geq 0$:

$$\|\bar{v}\|_{H_\beta^{k+2}(Q)} \leq C \left(\|\bar{f}\|_{H_\beta^k(Q)} + \sum_{m,l=0,1} \|\bar{G}_m^l\|_{H_\beta^{k+\frac{3}{2}-l}(Q)} \right) ,$$

which, together with (B.87), gives the desired estimate (B.86). □

For the original problem in Cartesian coordinates, equations (B.56a)-(B.56b) with boundary conditions (B.57), one can start from (B.82) with $k = 2$ and pass from polar to Cartesian coordinates via $v = A^{-1}\bar{v}$. Under such a transformation, the same inequality (B.86) follows for v, f and g_m^l as shown in [115]. The final step is to transfer the latter result from the first integral equations in terms of Ψ and Φ (B.55a) to the first order form (4.75), (4.76). Due to the fact that for $k \geq 0$:

$$\|u\|_{k+1,\beta} \leq \|\nabla^\perp \Psi\|_{k+1,\beta} + \|\nabla \Phi\|_{k+1,\beta} \leq \|\Psi\|_{k+2,\beta} + \|\Phi\|_{k+2,\beta} \leq \|v\|_{k+2,\beta} ,$$

the respective Theorem 4.15 for the first order system is obtained.

C Numerical method: further underpinnings

This appendix provides detailed convergence results for the benchmark cases of Sec. 5.3.1 from Ch. 5. Furthermore, the discretisation of the NS equations in terms of primitive variables is considered, being needed for comparison purposes; in particular the details of the classical scalar algebraic multigrid method are described as a background for Sec. 5.4.

C.1 Convergence results: Taylor-Couette flow

# variables	h	$\|u-u^*\|_1$	$\|u-u^*\|_0$	$\|u-u^*\|_\infty$
108	3.9018e-01	6.6642e+01	4.4920e+00	4.7164e+00
340	1.9603e-01	3.7467e+01	2.0548e+00	2.0976e+00
1188	9.8135e-02	1.9631e+01	6.1111e-01	5.8799e-01
4420	4.9082e-02	9.9522e+00	1.9213e-01	1.6371e-01
17028	2.4543e-02	4.9984e+00	5.3523e-02	4.5816e-02
66820	1.2272e-02	2.5067e+00	1.3965e-02	1.2563e-02
264708	6.1359e-03	1.2562e+00	3.7016e-03	3.4122e-03
1053700	3.0680e-03	6.2894e-01	1.0002e-03	9.2239e-04

Table C.1: Convergence results for test case [TF.1] with linear basis functions on 3-point triangular Lagrange elements and equal least-squares weighting of the four equations. The table shows the global errors in the energy norm, the $L^2(\Omega)$ norm and the maximum norm for increasingly refined grids.

# variables	h	$\|u_1-u_1^*\|_0$	$\|u_2-u_2^*\|_0$	$\|\phi_1-\phi_1^*\|_0$	$\|\phi_2-\phi_2^*\|_0$
108	3.9018e-01	1.0678e-01	1.0678e-01	4.1921e+00	1.6070e+00
340	1.9603e-01	1.0711e-01	1.0711e-01	1.9126e+00	7.3950e-01
1188	9.8135e-02	1.3206e-01	1.3206e-01	5.5079e-01	2.0474e-01
4420	4.9082e-02	7.5014e-02	7.5014e-02	1.5797e-01	5.4868e-02
17028	2.4543e-02	1.9400e-02	1.9400e-02	4.5230e-02	1.4900e-02
66820	1.2272e-02	3.4821e-03	3.4821e-03	1.2631e-02	4.0109e-03
264708	6.1359e-03	5.0244e-04	5.0244e-04	3.4833e-03	1.0782e-03
1053700	3.0680e-03	6.6205e-05	6.6205e-05	9.5349e-04	2.9001e-04

Table C.2: Convergence results for test case [TF.1]. The table shows for each variable the errors in the $L^2(\Omega)$ norm for increasingly refined grids.

# variables	h	$\|u_1 - u_1^*\|_\infty$	$\|u_2 - u_2^*\|_\infty$	$\|\phi_1 - \phi_1^*\|_\infty$	$\|\phi_2 - \phi_2^*\|_\infty$
108	3.9018e-01	1.5577e-01	1.5577e-01	4.7164e+00	2.3850e+00
340	1.9603e-01	1.5442e-01	1.5442e-01	2.0976e+00	1.1025e+00
1188	9.8135e-02	1.4155e-01	1.4155e-01	5.8799e-01	3.0920e-01
4420	4.9082e-02	9.5485e-02	9.5485e-02	1.6371e-01	8.2674e-02
17028	2.4543e-02	2.9861e-02	2.9861e-02	4.5816e-02	2.2081e-02
66820	1.2272e-02	5.9129e-03	5.9129e-03	1.2563e-02	5.8085e-03
264708	6.1359e-03	9.0827e-04	9.0827e-04	3.4122e-03	1.5215e-03
1053700	3.0680e-03	1.2442e-04	1.2442e-04	9.2239e-04	3.9863e-04

Table C.3: Convergence results for test case [TF.1]. The table shows for each variable the errors in the maximum norm for increasingly refined grids.

# variables	h	$\|u - u^*\|_1$	$\|u - u^*\|_0$	$\|u - u^*\|_\infty$
340	3.9018e-01	3.1112e-02	3.6462e-01	2.3545e-01
1188	1.9603e-01	7.5789e-03	3.6053e-02	2.3631e-02
4420	9.8135e-02	1.8745e-03	2.7477e-03	1.8377e-03
17028	4.9082e-02	4.6311e-04	2.1854e-04	1.4957e-04
66820	2.4543e-02	1.1481e-04	1.9497e-05	1.3523e-05
264708	1.2272e-02	2.8562e-05	1.9699e-06	1.3640e-06

Table C.4: Convergence results for test case [TF.2] with quadratic basis functions on 6-point triangular Lagrange elements and equal least-squares weighting of the four equations. The table shows the global errors in the energy norm, the $L^2(\Omega)$ norm and the maximum norm for increasingly refined grids.

# variables	h	$\|u_1 - u_1^*\|_0$	$\|u_2 - u_2^*\|_0$	$\|\phi_1 - \phi_1^*\|_0$	$\|\phi_2 - \phi_2^*\|_0$
340	3.9018e-01	1.3629e-02	1.3629e-02	1.3421e-01	2.3545e-01
1188	1.9603e-01	1.4659e-03	1.4659e-03	1.3115e-02	2.3631e-02
4420	9.8135e-02	1.0826e-04	1.0826e-04	1.0012e-03	1.8377e-03
17028	4.9082e-02	8.1272e-06	8.1272e-06	7.9191e-05	1.4957e-04
66820	2.4543e-02	8.0087e-07	8.0087e-07	6.9708e-06	1.3523e-05
264708	1.2272e-02	1.0634e-07	1.0634e-07	6.9379e-07	1.3640e-06

Table C.5: Convergence results for test case [TF.2] with quadratic basis functions on 6-point triangular Lagrange elements and equal least-squares weighting of the four equations. The table shows for each variable the errors in the $L^2(\Omega)$ norm for increasingly refined grids.

# variables	h	$\|u_1 - u_1^*\|_\infty$	$\|u_2 - u_2^*\|_\infty$	$\|\phi_1 - \phi_1^*\|_\infty$	$\|\phi_2 - \phi_2^*\|_\infty$
340	3.9018e-01	1.4513e-02	1.4513e-02	1.1859e-01	3.4439e-01
1188	1.9603e-01	1.3952e-03	1.3952e-03	1.1216e-02	3.4225e-02
4420	9.8135e-02	1.0378e-04	1.0378e-04	7.8973e-04	2.6288e-03
17028	4.9082e-02	8.5704e-06	8.5704e-06	5.8275e-05	2.1033e-04
66820	2.4543e-02	8.6220e-07	8.6220e-07	5.3609e-06	1.8709e-05
264708	1.2272e-02	9.9642e-08	9.9642e-08	6.1581e-07	1.8660e-06

Table C.6: Convergence results for test case [TF.2] with quadratic basis functions on 6-point triangular Lagrange elements and equal least-squares weighting of the four equations. The table shows for each variable the errors in the maximum norm for increasingly refined grids.

# variables	h	$\|u - u^*\|_1$	$\|u - u^*\|_0$	$\|u - u^*\|_\infty$
132	5.0000e-01	1.0992e+02	9.2902e-01	4.7789e-01
420	2.5000e-01	6.7602e+01	4.5945e-01	2.3316e-01
1476	1.2500e-01	3.6539e+01	2.7409e-01	1.6850e-01
5508	6.2500e-02	1.8712e+01	1.9639e-01	1.2867e-01
21252	3.1250e-02	9.4086e+00	6.4963e-02	5.0974e-02
83460	1.5625e-02	4.7170e+00	1.3868e-02	1.2551e-02
330756	7.8125e-03	2.3642e+00	2.2619e-03	2.1634e-03
1316868	3.9063e-03	1.1842e+00	3.2044e-04	3.1050e-04

Table C.7: Convergence results for test case [TF.3] with linear basis functions on 3-point triangular Lagrange elements and equal least-squares weighting of the four equations. The table shows the global errors in the energy norm, the $L^2(\Omega)$ norm and the maximum norm for increasingly refined grids.

# variables	h	$\|u - u^*\|_1$	$\|u - u^*\|_0$	$\|u - u^*\|_\infty$
420	5.0000e-01	9.8125e+00	1.6072e+00	8.4174e-01
1476	2.5000e-01	2.5304e+00	1.7601e-01	9.7660e-02
5508	1.2500e-01	6.2993e-01	1.4133e-02	9.3039e-03
21252	6.2500e-02	1.5262e-01	1.0850e-03	7.6415e-04
83460	3.1250e-02	3.6454e-02	9.6779e-05	9.3713e-05
330756	1.5625e-02	8.8204e-03	7.8854e-06	7.7389e-06

Table C.8: Convergence results for test case [TF.3] with quadratic basis functions. The table shows the global errors in the energy norm, the $L^2(\Omega)$ norm and the maximum norm for increasingly refined grids.

C.2 Convergence results: Colliding flow

# variables	h	CF.1	CF.2	CF.3	CF.4
100	5.0000e-01	1.4966e+00	3.3853e+00	2.8345e-01	3.1376e-01
324	2.5000e-01	2.4242e-01	1.6120e-01	3.2931e-02	3.5283e-02
1156	1.2500e-01	3.5569e-02	1.3945e-02	3.9720e-03	5.4919e-03
4356	6.2500e-02	4.8052e-03	1.1789e-03	4.9275e-04	7.9649e-04
16900	3.1250e-02	6.2595e-04	1.0535e-04	6.1552e-05	9.3813e-05
66564	1.5625e-02	7.9971e-05	1.0413e-05	7.6960e-06	2.2858e-05
264196	7.8125e-03	1.0110e-05	1.1474e-06	9.6205e-07	4.0205e-06

Table C.9: $L^2(\Omega)$-norm of the solution error ($\|u-u^*\|_0$) for cases [CF.1]-[CF.4] of colliding flow and for increasingly refined grids. The speed of convergence is comparable for all types of boundary conditions and in each case the optimal cubic rate for second-order FEM basis functions is reached, that is $\|u - u^*\|_0 \sim ch^3$.

# variables	h	CF.1	CF.2	CF.3	CF.4
100	5.0000e-01	1.1073e+04	2.1678e+04	6.4855e+02	3.2408e+02
324	2.5000e-01	4.1697e+04	3.2151e+04	2.7153e+03	3.2180e+03
1156	1.2500e-01	1.7898e+05	7.8765e+04	7.1439e+03	2.1099e+04
4356	6.2500e-02	7.7416e+05	1.9773e+05	2.1354e+04	2.0543e+05
16900	3.1250e-02	3.3629e+06	5.0759e+05	8.2021e+04	2.6761e+06
66564	1.5625e-02	1.4609e+07	1.1090e+06	2.9502e+05	5.0867e+07
264196	7.8125e-03	6.3289e+07	4.1829e+06	1.1459e+06	1.1974e+09

Table C.10: 2-norm condition numbers of the system matrices resulting from test cases [CF.1]-[CF.4] of colliding flow and for increasingly refined grids. The condition numbers grow proportional to h^2 apart for case CF.4.

C.3 Taylor-Hood discretisation of the NS equations

The Stokes equations describe the flow of an incompressible viscous fluid under the assumption that inertial effects can be neglected in comparison to inner friction. The two-dimensional variant of the equations without external forces and with pure Dirichlet boundaries is:

$$\eta\Delta u - \nabla p = 0 \quad \text{in } \Omega\,, \tag{C.1}$$

$$\operatorname{div}(u) = 0 \quad \text{in } \Omega\,, \tag{C.2}$$

$$u = g \quad \text{auf } \partial\Omega\,. \tag{C.3}$$

Here, $u = (u_x, u_y)$ denotes the velocity vector, p the pressure and η the dynamic viscosity. The domain and its boundary is denoted by Ω and $\partial\Omega$. Equations (C.1), (C.2) result from the Lagrangian density:

$$L(u_i, \partial_j u_i, p) = 2\eta \sum_{i,j=1}^{2} D_{ij} D_{ij} - p \sum_{i=1}^{2} \frac{\partial u_i}{\partial x_i}\,, \tag{C.4}$$

with strain rate tensor:

$$D_{ij} = \frac{1}{2}\left[\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right],$$

so that the variational principle behind (C.4) can be interpreted as a minimisation of dissipation. The pressure p plays the role of a Lagrangian multiplier necessary in order to fulfil the incompressibility condition (C.2), which is therefore obtained by variation with respect to p. The two equations (C.1) follow by variation with respect to u_x and u_y.

In the following let $L_2(\Omega)$ denote the Lebesgue space of quadratic integrable functions:

$$L_2(\Omega) := \left\{ f : \Omega \to \mathbb{R} \text{ measurable}, \int_\Omega |f|^2 \,\mathrm{d}\Omega < \infty \right\},$$

and $H^1(\Omega)$ the standard Sobolev space of real-valued functions $g \in L_2(\Omega)$ with all first order partial derivatives existing and being in $L_2(\Omega)$. H_0^1 denotes the Sobolev space H^1 with zero boundary conditions. Now, define the bilinear form:

$$a(v, u) = \eta \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}\Omega\,, \tag{C.5}$$

$$b_1(v, p) = -\int_\Omega p \cdot \frac{\partial v}{\partial x} \,\mathrm{d}\Omega\,, \quad b_2(v, p) = -\int_\Omega p \cdot \frac{\partial v}{\partial y} \,\mathrm{d}\Omega\,, \tag{C.6}$$

and the function spaces:

$$V_0 = H_0^1(\Omega)\,, \quad V = \left\{ v \in [H^1(\Omega)]^2 \;\middle|\; v = g \text{ auf } \partial\Omega \right\},$$
$$W = \left\{ w \in L_2(\Omega) \;\middle|\; \int_\Omega w \,\mathrm{d}\Omega = 0 \right\},$$

then a weak formulation of the Stokes system (C.1) – (C.3) is given by:

$$\text{Find } (u_x, u_y) \in V,\ p \in W, \text{ such that:} \tag{C.7}$$

$$a(v_x, u_x) + b_1(v_x, p) = 0 \quad \text{for all } v_x \in V_0\,, \tag{C.8}$$

$$a(v_y, u_y) + b_2(v_y, p) = 0 \quad \text{for all } v_y \in V_0\,, \tag{C.9}$$

$$b_1(u_x, w) + b_2(u_y, w) = 0 \quad \text{for all } w \in W\,. \tag{C.10}$$

In practice, a frequently used choice of mixed FE spaces fulfilling the so-called Inf-Sup

compatibility condition employs Lagrangian elements with different polynomial degree: k-th order polynomials for the velocities and $(k-1)$-th order polynomials for the pressure field. This combination is called the Taylor-Hood element and known to be stable for $k \geq 2$. The Taylor-Hood FEM calculations for a comparative study with the LSFEM in Sec. 5.3.2 are based on $(\mathcal{P}_2, \mathcal{P}_1)$ elements for an irregular triangulation; velocity and pressure variables are distributed on the grid nodes as shown in Fig. C.1.

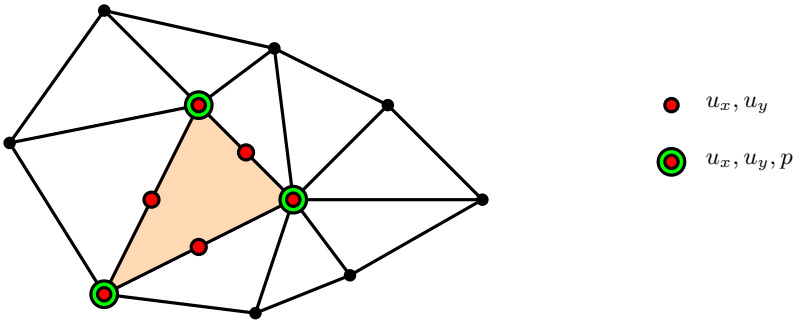


Figure C.1: Variable distribution on a FE triangulation: the pressure is discretised at corner points only (linear approximation), the velocities additionally at the edge centres (quadratic approximation) also.

Consider Fig. C.1, then for each node of the given grid quadratic basis functions v_i can be found and for each corner node linear basis functions w_i forming the discrete finite element subspaces $V^h \subset V$, $W^h \subset W$:

$$V^h = \{v_1, v_2, \ldots, v_{n_u}\}\,, \quad W^h = \{w_1, w_2, \ldots, w_{n_p}\}\,, \quad n_u > n_p\,;$$

the unknown velocity and pressure fields are approximated by the basis functions according to:

$$u_x = \sum_{i=1}^{n_u} c_i^1 v_i\,, \quad u_y = \sum_{i=1}^{n_u} c_i^2 v_i\,, \quad p = \sum_{i=1}^{n_p} c_i^3 w_i\,.$$

Thus the discret version of (C.7) – (C.10) transforms to a linear system of equations: with matrices $A \in \mathbb{R}^{n_u \times n_u}$, $B_1, B_2 \in \mathbb{R}^{n_u \times n_p}$:

$$\begin{bmatrix} A & 0 & B_1 \\ 0 & A & B_2 \\ B_1^T & B_2^T & 0 \end{bmatrix} \cdot \begin{bmatrix} c^1 \\ c^2 \\ c^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}\,, \tag{C.11}$$

in which:

$$A_{ij} = a(v_i, v_j)\,, \quad (B_1)_{ij} = b_1(v_i, w_j)\,, \quad (B_2)_{ij} = b_2(v_i, w_j)\,. \tag{C.12}$$

The Dirichlet boundary conditions are incorporated subsequently and the periodic boundary conditions are automatically fulfilled by a fictitious linkage of the grid points on the left and right-hand side of the domain boundary.

In order to solve the full Navier-Stokes equations with non-linear convection terms a Newton linearisation is applied such that for $n = 1, 2, \ldots$ an iteration process of the form:

$$\begin{aligned}
\rho\left[u_x^{(n-1)}\frac{\partial u_x^{(n)}}{\partial x} + u_x^{(n)}\frac{\partial u_x^{(n-1)}}{\partial x} + u_y^{(n-1)}\frac{\partial u_x^{(n)}}{\partial y} + u_y^{(n)}\frac{\partial u_x^{(n-1)}}{\partial y}\right] + \frac{\partial p}{\partial x} - \eta\Delta u_x = \\
\rho\left[u_x^{(n-1)}\frac{\partial u_x^{(n-1)}}{\partial x} + u_y^{(n-1)}\frac{\partial u_x^{(n-1)}}{\partial y}\right], \\
\rho\left[u_x^{(n-1)}\frac{\partial u_y^{(n)}}{\partial x} + u_x^{(n)}\frac{\partial u_y^{(n-1)}}{\partial x} + u_y^{(n-1)}\frac{\partial u_y^{(n)}}{\partial y} + u_y^{(n)}\frac{\partial u_y^{(n-1)}}{\partial y}\right] + \frac{\partial p}{\partial y} - \eta\Delta u_y = \\
\rho\left[u_x^{(n-1)}\frac{\partial u_y^{(n-1)}}{\partial x} + u_y^{(n-1)}\frac{\partial u_y^{(n-1)}}{\partial y}\right], \\
\frac{\partial u_x^{(n)}}{\partial x} + \frac{\partial u_y^{(n)}}{\partial y} = 0\,,
\end{aligned}$$

is obtained. As a starting point, $n = 0$, the linear Stokes solution is employed. If the bilinear form is defined similar to the above, i.e.:

$$c(v, u, u_1, u_2) = \rho\int_\Omega v\left[u_1\frac{\partial u}{\partial x} + u_2\frac{\partial u}{\partial y}\right]\mathrm{d}\Omega\,, \tag{C.13}$$

$$d_1(v, u, u_1) = \rho\int_\Omega v\left[u\frac{\partial u_1}{\partial x}\right]\mathrm{d}\Omega\,, \quad d_2(v, u, u_1) = \rho\int_\Omega v\left[u\frac{\partial u_1}{\partial y}\right]\mathrm{d}\Omega\,, \tag{C.14}$$

the method can be described by a sequence of linear systems:

$$\begin{bmatrix} A + C_1^{(n-1)} & D_1^{(n-1)} & B_1 \\ D_2^{(n-1)} & A + C_2^{(n-1)} & B_2 \\ B_1^T & B_2^T & 0 \end{bmatrix} \cdot \begin{bmatrix} c^1 \\ c^2 \\ c^3 \end{bmatrix} = \begin{bmatrix} E_1^{(n-1)} \\ E_2^{(n-1)} \\ 0 \end{bmatrix}\,, \tag{C.15}$$

with matrices $C_1, C_2, D_1, D_2 \in \mathbb{R}^{n_u \times n_u}$ and vectors $E_1, E_2 \in \mathbb{R}^{n_u \times 1}$:

$$(C_1)_{ij} = c(v_i, v_j, u_x^{(n-1)}, u_y^{(n-1)}) + d_1(v_i, v_j, u_x^{(n-1)})\,, \tag{C.16}$$

$$(C_2)_{ij} = c(v_i, v_j, u_x^{(n-1)}, u_y^{(n-1)}) + d_2(v_i, v_j, u_y^{(n-1)})\,, \tag{C.17}$$

$$(D_1)_{ij} = d_2(v_i, v_j, u_x^{(n-1)})\,, \qquad (D_2)_{ij} = d_1(v_i, v_j, u_y^{(n-1)})\,, \tag{C.18}$$

$$(E_1)_i = c(v_i, u_x^{(n-1)}, u_x^{(n-1)}, u_y^{(n-1)})\,, \qquad (E_2)_i = c(v_i, u_y^{(n-1)}, u_x^{(n-1)}, u_y^{(n-1)})\,. \tag{C.19}$$

As a stopping criterion a residual of $\mathrm{Res} < 10^{-6}$ is fixed. Subsequent to the solution for the velocity components and the pressure, the velocity derivatives which are crucial to the

calculation of the wall shear stress along the upper plate are continuously reconstructed by standard methods.

C.4 Scalar algebraic multigrid: the details

This section is attached to the basic considerations concerning scalar AMG as utilised in Chapter 5, Sec. 5.4. Here a few more details are provided which are mandatory for the understanding of the AMG strategy developed for the linear systems resulting from the first integral LSFEM. In particular, the meaning of algebraic smoothing is clarified and standard coarsening and interpolation methods summarised.

C.4.1 Algebraic smoothing

Smoothing and coarse-grid correction play the same role in both AMG and GMG but with algebraic smoothness having a different meaning to geometric smoothness. In a geometric context the term "smooth" is used in a natural and restrictive way, i.e., an error is called geometrically smooth if it *can* be well approximated on a predefined coarser grid; thus the smoothing property in GMG is always related to two specified grids; conversely in AMG no predefined grids exist and the classical definition is meaningless; instead, an error is said to be algebraically smooth if it converges slowly with regard to a smoothing operator S, i.e. $Se \approx e$, or in other words if it is not efficiently reduced by the smoothing operator and thus *must* be approximated on a coarser level.

Before proceeding further, additional notation is introduced for the description of algebraic smoothness. A large and sparse n-dimensional linear system with a symmetric, positive definite matrix A is assumed, the latter property indicated by $A > 0$. Related to the matrix A, three scalar products are defined for $u, v \in \mathbb{R}^n$:

$$(u,v)_0 = (D_h u, v)_E, \quad (u,v)_1 = (A_h u, v)_E, \quad (u,v)_2 = (D_h^{-1} A_h u, A_h v)_E\,, \tag{C.20}$$

with the corresponding norms $\|.\|_i$ $(i = 0, 1, 2)$[1]. Here, $(.,.)_E$ denotes the Euclidean scalar product and $D_h = \mathrm{diag}(A_h)$. $(.,.)_1$ and $\|.\|_1$ are also called the energy product and energy norm. The spectral radius of a matrix is denoted by $\rho(.)$ as usual.

(a) Eigenvectors and Eigenvalues

The eigenvectors ϕ of $D^{-1}A$ play a significant role in the context of algebraic smoothness. The eigenvectors that belong to the smallest eigenvalues λ normally lead to the slowest

[1]The discrete inner products and norms given by (C.20) are valid throughout Sec. C.4 only and are not to be confused with the similar notation for Sobolev norms used elsewhere.

convergence of the relaxation process and thus correspond to the above definition of algebraically smooth error components. This can be easily verified for a ω-Jacobi iteration with suitable under-relaxation. Small eigenvalues λ of $D^{-1}A$ correspond to eigenvalues of the ω-Jacobi operator $S = (I - \omega D^{-1}A)$ which are close to one. This is similar to related linear iteration methods, for example GS relaxation.

Next, some basic correlations between the above defined norms belonging to the scalar products (C.20) are discussed [242]:

Lemma C.1. *Let $A > 0 \in \mathbb{R}^{n\times n}$. Then for all vectors $e \in \mathbb{R}^n$:*

$$\|e\|_1^2 \leq \|e\|_0\|e\|_2, \quad \|e\|_2^2 \leq \rho(D^{-1}A)\|e\|_1^2, \quad \|e\|_1^2 \leq \rho(D^{-1}A)\|e\|_0^2\,;$$

and for the eigenvectors of $D^{-1}A$:

$$D^{-1}A\phi = \lambda\phi \quad \Rightarrow \quad \|\phi\|_2^2 = \lambda\|\phi\|_1^2 \quad \textit{and} \quad \|\phi\|_1^2 = \lambda\|\phi\|_0^2\,. \tag{C.21}$$

The eigenvalues and eigenvectors of $D^{-1}A$ in combination with the result (C.21) demonstrate the significance of the choice of norms in the context of smoothing: if (C.21) is applied to an algebraically smooth error component $e = \phi$ the corresponding eigenvalue λ is close to one and therefore the norms have considerably differing magnitude:

$$\|\phi\|_2 \ll \|\phi\|_1 \quad \text{and} \quad \|\phi\|_1 \ll \|\phi\|_0\,.$$

Conversely, norms of algebraically non-smooth errors are of comparable magnitude, leading to the conclusion that an algebraically smooth error can be identified simply by comparing different norms. This yields the following formal definition of the smoothing property:

Definition C.2. *The relaxation operator S will possess the required smoothing property with regard to the matrix $A > 0$ if:*

$$\|Se\|_1^2 \leq \|e\|_1^2 - \sigma\|e\|_2^2 \quad (\sigma > 0)\,, \tag{C.22}$$

holds for all e and σ independent of e.

The relaxation operator S efficiently reduces the error e provided $\|e\|_2$ is large relative to $\|e\|_1$. In contrast, relaxation operations become inefficient if $\|e\|_2 \ll \|e\|_1$; the error is then called algebraically smooth. Furthermore, S is said to fulfil the smoothing property with regard to a class of matrices $\mathcal{A}$ if S satisfies inequality (C.22) uniformly for all $A \in \mathcal{A}$ with an identical σ. A necessary requirement for (C.22) to hold is $\sigma\|e\|_2^2 \leq \|e\|_1^2$ which is equivalent to $\rho(D^{-1}A) \leq 1/\sigma$. Consequently, for fulfilment of the smoothing property in $\mathcal{A}$, the spectral radius $\rho(D^{-1}A)$ has to be uniformly bounded from below for all $A \in \mathcal{A}$.

(b) Smoothing methods

Classical relaxation schemes such as Jacobi or GS are usually employed as smoothers; the computational effort is comparably low and the smoothing effect sufficient for typical classes of matrices. Such a typical class in the context of elliptic differential equations is that of symmetric and positive definite M-Matrices the row sum of which is close to zero. The M-matrix property is defined as follows:

Definition C.3. *A matrix $A \in \mathbb{R}^{N\times N}$ is called a M-Matrix, if the following conditions are fulfilled:*

(i) $a_{ii} > 0$ *for* $i = 1, \ldots, N$,

(ii) $a_{ij} \leq 0$ *for all* $i \neq j$,

(iii) A^{-1} *exists and it holds* $A^{-1} \geq 0$.

A well-known result concerning the smoothing property of the GS method, which is similarly valid for Jacobi relaxation, is stated by Stüben [242], namely:

Proposition C.4. *For all symmetric M-matrices Gauss-Seidel relaxation fulfils the smoothing property* (C.22) *uniformly with* $\sigma = 1/4$. *This similarly holds for all matrices A which are obtained from symmetric M-matrices by symmetric permutation of arbitrarily many signs.*

In addition, many classical ILU schemes [211] which are frequently used for preconditioning Krylov subspace iterations exhibit a strong smoothing effect on the error. The computation of an ILU decomposition requires a lot more computational effort in comparison to point-wise relaxation schemes, but, due to the higher number of variables involved in the smoothing process, may provide a significantly higher quality of smoothing.

(c) Interpretation of algebraically smooth errors

For a better understanding of coarsening and interpolation further discussion is required of the meaning of algebraically smooth errors. As noted in Sec. 5.4, an algebraic smooth error e is characterised by the property $Se \approx e$ according to the above considerations implying $\|e\|_2 \ll \|e\|_1$. Expressing this in terms of the residual $r = Ae$ gives:

$$(D^{-1}r, r)_E \ll (e, r)_E.$$

Therefore algebraically smooth errors are on average characterised by scaled residuals which are significantly smaller than the error itself leading to the heuristic assumption:

$$|r_i| \ll a_{ii}|e_i| \,,$$

for row indices i of an algebraically smooth error e and its corresponding residual r. This again yields the approximate relationship:

$$a_{ii}e_i \approx -\sum_{j\in N_i} a_{ij}e_j\,, \tag{C.23}$$

in which $N_i := \{j \in \Omega : j \neq i, a_{ij} \neq 0\}$ denotes the index set of the neighbouring variables of i. Thus, although the error might be large globally, e_i may be approximated well by a function in terms of the neighbouring errors e_j.

(d) Algebraically smooth errors and strong couplings

In the following the relationship between the direction of algebraic smoothness and the variable coupling in the system matrix is identified. An algebraic smooth error e satisfying $\|e\|_2 \ll \|e\|_1$ implies $\|e\|_1 \ll \|e\|_0$. The latter, for symmetric matrices A, is equivalent to:

$$\frac{1}{2}\sum_{i,j}(-a_{ij})(e_i - e_j)^2 + \sum_i s_i e_i^2 \ll \sum_i a_{ii}e_i^2\,, \tag{C.24}$$

with the sum of the i-th row of A denoted by $s_i = \sum_j a_{ij}$. Moreover, for M-matrices $a_{ij} \leq 0$ $(j \neq i)$ holds and if additionally the row sums are close to zero, that is $s_i \approx 0$, for each i the following estimate:

$$\sum_{j\neq i} \frac{|a_{ij}|}{a_{ii}} \frac{(e_i - e_j)^2}{e_i^2} \ll 1\,,$$

can be derived. This motivates the following statement which is firstly confined to symmetric M-matrices with approximately zero row sums:

Lemma C.5. *An algebraic smooth error varies slowly in directions of strong negative couplings, that is the error from e_i to e_j will change only marginally if $|a_{ij}|/a_{ii}$ is relatively large. Relaxation methods that satisfy the smoothing property will smooth the error in directions of strong negative couplings.*

The above results can be generalised to certain classes of matrices with positive non-diagonal entries [242] but it is not necessary to consider these further in the context of the present work.

C.4.2 Coarsening procedure

Consider as in Sec. 5.4 a two-level arrangement where a suitable disjoint splitting of the fine-level variables $\Omega = C \cup F$ into coarse-level variables (C) and the remaining variables (F) is required. In order to obtain fast convergence, the algebraically smooth error should

be well approximated by interpolation and this in turn is best achieved the stronger the coupling between the F and C-variables is. As the overall efficiency of the method depends on both the convergence rate and the computational effort per cycle, the coarsening process should be designed so as to minimise the number of C-variables while retaining a sufficiently strong coupling between the F and C-variables. Furthermore, the splitting into F and C-variables should preferably be performed in a homogeneous manner in the sense that F-variables are surrounded by C-variables from which they can be interpolated. This requirement usually improves the interpolation considerably leading to faster convergence; such a simple splitting scheme is given by so-called standard coarsening, see (a) below.

The imperative of a strong coupling between F- and C-variables does not necessarily mean that this coupling has to exist in a direct manner; not every F-variable has to be directly strongly-coupled to a C-variable but the link can exist indirectly via a coupling to other F-variables which are again strongly coupled to a C-variable. This approach allows for a distinct reduction in the number of necessary C-variables and opens up the prospect of a drastic reduction in computation time which often compensates for the disadvantage of a limited smoothing and interpolation quality in practice. Such an aggressive coarsening strategy is presented in (b).

(a) Standard coarsening

The term *standard coarsening* comprises splitting schemes, based on direct couplings between C- and F-variables. According to a pre-determined requirement, for each F-variable i a minimal number of direct and sufficiently strong couplings $j \in N_i$ to the set of coarse level variables has to be identified. For the applications considered in the present work, the strong couplings are mostly negative and the splitting algorithm can be confined to those negative couplings. In other words coarsening takes place in the direction of slowly varying algebraically smooth errors, see Lemma C.5. Here, a specific variable i is said to have a strong negative coupling to another variable j if:

$$-a_{ij} \geq \varepsilon_{str} \max_{a_{ik}<0} |a_{ik}| \quad \text{with fixed} \quad 0 < \varepsilon_{str} < 1\,. \tag{C.25}$$

The set of strong couplings of the variable i is denoted by:

$$S_i = \{j \in N_i : i \text{ strongly coupled to } j\}\,. \tag{C.26}$$

All other couplings, including strong positive ones, are treated as weak. A standard value for this coupling constant is $\varepsilon_{str} = 0.25$ [242]. The relationship of strong couplings will normally be non-symmetric even if the matrix A is symmetric; therefore the set of transposed couplings to a variable i must be defined also, that is the set of variables

coupled to i:

$$S_i^T = \{j \in \Omega : i \in S_j\}\,.$$

A simple coarsening procedure works as follows: An arbitrary first variable i is chosen to be a C-variable and all variables $j \in S_i^T$ strongly coupled to i are classified as F-variables. The rest of the non-classified variables are treated in the same way until all variables are either in the fine-level or the coarse-level set. In order to obtain a uniform distribution of C- and F-variables the respective next C-variable is not chosen randomly but according to a specified priority order. For this purpose a priority measure λ_i is defined indicating the urgency for i to be chosen as the next C-variable:

$$\lambda_i = |S_i^T \cap U| + 2|S_i^T \cap F| \quad (i \in U)\,. \tag{C.27}$$

Here, U is the set of currently unspecified variables and $|\cdot|$ denotes the number of elements in the specified set. λ_i indicates the value of i as a C-variable. At the beginning variables are preferred to which many other variables are strongly coupled; later on variables are preferred to which many F-variables are strongly coupled. Figs. C.2 and C.3 illustrate this process.

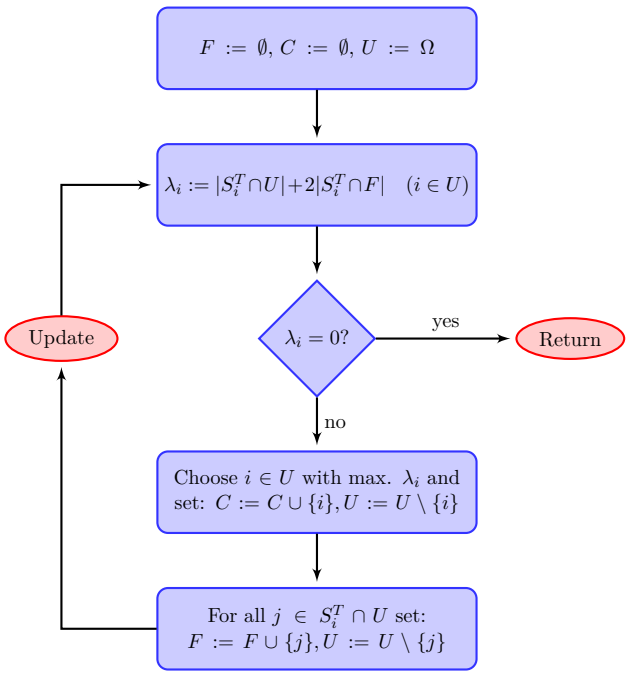


Figure C.2: Standard coarsening algorithm schematic

In a preceding step of the algorithm variables without couplings, or which belong to strongly diagonally dominant rows, are filtered and marked as F-variables; these variables do not need to be interpolated. At the end of the algorithm all F-variables, except for the filtered ones, have at least one strong coupling to a C-variable but it may be the case that some U-variables are left without any strong coupling to a C-variable. Neither of these variables have strong couplings between them nor is any F-variable strongly coupled to them. However, each of these variables is at least strongly coupled to at least one F-variable; that means they can be marked as F-variables and will be interpolated indirectly via their strong F-couplings (Sec. C.4.3). Finally, it is noted that the resulting set of C-variables is close to the maximum set of variables not being coupled to each other (maximum independent set).

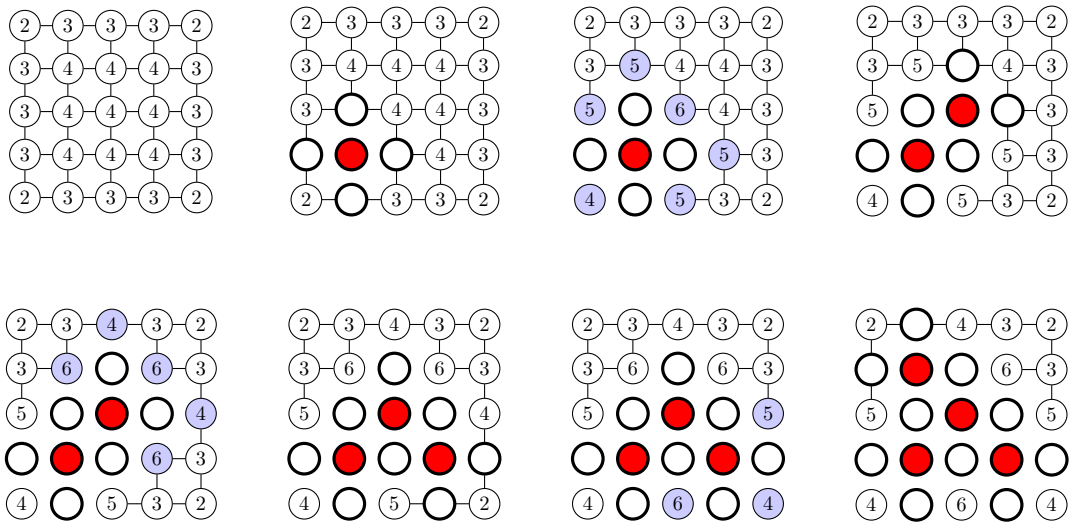


Figure C.3: A schematic illustration of standard coarsening for a five-point finite difference stencil. At the beginning (top-left) every U-variable is initialised with a number representing of its strong coupling to its immediate neighbours (here it is assumed that any delineated connection represents a strong coupling). Proceeding chronologically from top-left to bottom-right, one of the variables with highest priority is chosen as a C-variable (red), the neighbours become F-variables (white), then for each F-variable the priority of the neighbouring U-variables (blue) is increased by one. Next, a new C-variable is chosen and so on.

(b) Aggressive coarsening

In many applications the discretisation involves relatively small difference stencils and in those cases a coarsening procedure based on direct couplings requires many C-variables, thus resulting in a Galerkin operator with a high level of complexity. This becomes

noticeable especially in the first coarsening step which in turn contributes significantly to the overall complexity of the setup phase.

In order to achieve a more aggressive coarsening, indirect couplings between variables are considered. A variable i is said to be coupled to a variable j over a path of length l if a sequence of variables $i_0, i_1, ..., i_l$ exists with $i = i_0$ and $j = i_l$ such that $i_{k+1} \in S_{i_k}$ for $k = 0, 1, 2, ..., l-1$. Let $p \geq 1$ and $l \geq 1$, then i is said to be strongly (p, l)-coupled to j if at least p paths with a length of at most l exist providing a strong coupling between i and j. Such an aggressive coarsening is obtained if in the algorithm of Fig. C.2 for given p and l the set S_i is replaced by:

$$S_i^{p,l} = \{j \in \Omega : i \text{ strongly } (p,l)\text{-coupled to } j\}\,.$$

In practice only strategies involving very low values of p and l pay off, therefore coarsening with $S_i^{1,2}$ and $S_i^{2,2}$ is also seperately called A1 and A2 coarsening, respectively. Normally rarely more than one aggressive coarsening per setup is used and is most often performed directly at the beginning of the coarsening procedure.

If the coarsening algorithm is applied directly to $S_i^{p,l}$ rather than S_i the complete coupling information and information about $(S_i^{p,l})^T$ have to be saved for all variables i. The cost for this can be reduced by employing a two-step algorithm: in the first step the algorithm of Fig. C.2 is performed in its original form as explained in Sec. 2.2.1.; in the second step strong couplings are solely defined between C-variables (via neighbouring F-variables):

$$\hat{S}_i^{p,l} = \{j \in C : i \text{ strongly } (p,l)\text{-coupled to } j\}\,.$$

Now, the standard coarsening algorithm is applied once again but confined to the C-variables and the resulting variable set defines the next coarser level (see Fig. C.4).

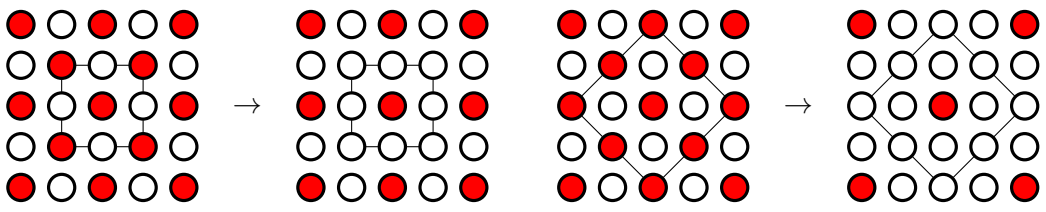


Figure C.4: Schematic illustration of two types of aggressive coarsening, A2 (left) and A1 (right), for the case of a five-point stencil. The marked boxes in each case emphasise the range of the strong couplings in the sense of $\hat{S}_i^{2,2}$ and $\hat{S}_i^{1,2}$.

C.4.3 Interpolation

Having defined the split into C and F-variables on each grid level, it is now necessary to specify the corresponding different interpolation operators. These operators in turn depend on whether standard or aggressive coarsening has been employed. In the first case direct (Par. (a)) or standard interpolation (Par. (b)) is used, in the second case multi-pass interpolation (Par. (c)). Both forms of interpolation can be improved by additional relaxation steps leading to, a so-called, Jacobi interpolation [242] but which is not considered here. However, in contrast to direct interpolation the standard form already involves indirect strong couplings, i.e. an extension of the interpolation radius is effected, which can necessitate a truncation of the interpolation operator complexity before eventually computing the Galerkin operator.

In what follows additional conventions are required. Let S_i be defined as in (C.26) and furthermore let:

$$C_i = C \cap N_i, \quad C_i^s = C \cap S_i$$
$$F_i = F \cap N_i, \quad F_i^s = F \cap S_i \,.$$

(a) Direct interpolation

Sec. C.4.1(d) revealed that for symmetric M-matrices the algebraically smooth error varies slowly in directions of strong coupling, that is the error in a variable i can be well approximated by a weighted average error of the strongly coupled neighbouring variables. For each $i \in F$ a set of interpolation variables $P_i = C_i^s$ is defined and (C.23) is approximated by:

$$a_{ii}e_i + \sum_{j \in N_i} a_{ij}e_j = 0 \quad \Rightarrow \quad a_{ii}e_i + \alpha_i \sum_{k \in P_i} a_{ik}e_k \,, \tag{C.28}$$

with:

$$\alpha_i = \frac{\sum_{j \in N_i} a_{ij}}{\sum_{k \in P_i} a_{ik}} \,.$$

This leads to the interpolation formula:

$$e_i = \sum_{k \in P_i} w_{ik}e_k \quad \text{mit} \quad w_{ik} = -\alpha_i a_{ik}/a_{ii} \quad (i \in F, k \in P_i) \,,$$

which only considers direct couplings, thus the descriptor *direct* interpolation, see also Fig. C.5.

The above procedure can be applied as long as $C_i^s \neq \emptyset$ but this condition is, when using standard coarsening, satisfied for all F-variables except for special cases; these always have

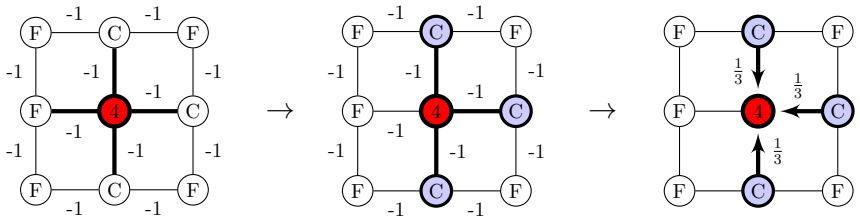


Figure C.5: A schematic illustration of direct interpolation, from left to right, for a five-point stencil. The middle F-variable (red) is strongly coupled to all four neighbours (diagonal entry: 4, all minor diagonal entries: -1) but is only interpolated from the neighbouring C-variables (blue). The interpolating weights, in this case are all equal to 1/3, depend on the magnitude of the matrix entries.

at least one strong coupling to a regular F-variable and thus can be interpolated indirectly as shown in the following.

(b) Standard interpolation

As in Sec. C.4.3(a) a predetermined C/F-splitting according to the standard coarsening algorithm is assumed ensuring a strong coupling between F and C-variables. Nevertheless, this does not guarantee the requirement of the two-level theory (see Sec. C.4.4) that for each F-variable a certain percentage of the total coupling is represented in the coarse level set (defined by the parameter τ). Although this argument does not pose any problem in practice as the coarsening algorithm provides a sufficient F/C coupling, the problem can be easily overcome. The direct interpolation is modified such that for each variable $i \in F$ the strong coupling to other F-variables is indirectly incorporated as well. Instead of approximating the i-th equation on the left-hand side of (C.28) directly, all associated e_j with $j \in F_i^s$ are replaced by means of the j-th equation, that is:

$$e_j \to -\sum_{k \in N_j} a_{jk} e_k / a_{jj} \,, \tag{C.29}$$

is replaced for all $j \in F_i^s$ and a new equation is obtained for e_i:

$$\hat{a}_{ii} e_i + \sum_{j \in N_i} \hat{a}_{ij} e_j = 0 \quad \text{mit} \quad \hat{N}_i = \{ j \neq i : \hat{a}_{ij} \neq 0 \} \,. \tag{C.30}$$

If now P_i is redefined as $P_i = C_i^s \cup \bigcup_{j \in F_i^s} C_j^s$ and all a are replaced by $\hat{a}$ and N_i by $\hat{N}_i$, respectively, the interpolation can be performed exactly as in Sec. 2.4.1, see Fig. C.6. This modification can improve the quality of direct interpolation significantly as the approximation from (C.30) reduces the error in comparison to (C.28).

Direct and standard interpolation can also be combined in a complementary way. For

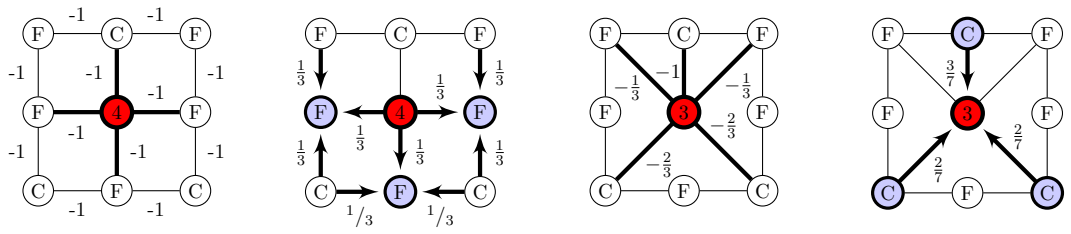


Figure C.6: standard interpolation for the five-point stencil. The middle F-variable (red) is strongly coupled to all four neighbours (diagonal entry: 4, all minor diagonal entries: -1). The error of the neighbouring F-variables (blue) is eliminated according to (C.29) and new matrix entries (C.30) are generated. The error of the middle F-variable is now directly interpolated by the new and old neighbouring C-variables.

example one can basically use direct interpolation together with standard interpolation for those variables for which the F/C-coupling appears too weak. Additionally, the set of interpolation variables could be extended by also taking into account the weak couplings.

Standard interpolation may result in large sets of interpolation variables P_i such that the complexity of the Galerkin operator significantly increases in the direction of coarsening; a drawback which can be controlled by a truncation of interpolation operators. The interpolation weights of variables being far away from i are usually considerably smaller than the largest ones. Thus the interpolation operators should be truncated before computing the coarse level Galerkin operator; for example, by ignoring weights which are smaller by a factor ε_{tr} than the maximum weight. In practice $\varepsilon_{tr} = 0.2$ is a commonly used value.

(c) Multipass interpolation

Multipass interpolation can be utilised in the context of aggressive coarsening strategies. Here, if possible, direct interpolation is performed and the remaining variables are interpolated via strongly coupled F-variables. The procedure is as follows (see also Fig. C.7):

1. For all variables $i \in F$ with $C_i^s \neq \emptyset$ use direct interpolation and collect the indices of these variables in F^*. If $F = F^* \rightarrow$ stop, else continue.

2. For all $i \in F \setminus F^*$ with $S_i \cap F^* \neq \emptyset$ proceed as follows: Take the i-th equation (left-hand side in (C.28)) replacing e_j for all $j \in S_i \cap F^*$ by:

$$e_j \rightarrow \sum_{k \in P_j} w_{jk} e_k \,,$$

leading to a modified form of equation (C.30) for e_i. Define the set of interpolation variables for i as $P_i = \bigcup_{j \in S_i \cap F^*} P_j$; an interpolation formula analogous to standard

interpolation can then be derived. After having treated all variables, the particular ones assigned an interpolation formula are added to F^*.

3. If $F = F^* \rightarrow$ stop, else go to step 2.

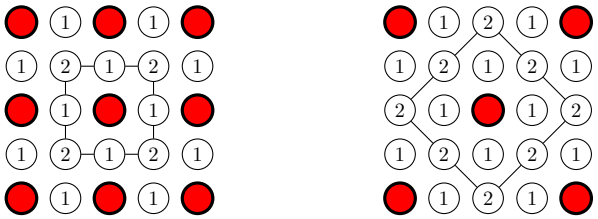


Figure C.7: A schematic illustration of multipass interpolation for a five-point stencil using aggressive A2 (left) and A1 (right) coarsening, respectively. The numbers indicate in which pass the respective F-variable has been reached.

C.4.4 Convergence of the two-level cycle

In the following subsections a few key results associated with the convergence theory of AMG are summarised.

(a) The variational principle for positive definite problems

The notation of the two-level cycle from Sec. 5.4.1 is adopted and a symmetric, positive definite system matrix assumed. With regard to the transfer operators let $I_h^H = (I_H^h)^T$ such that the Galerkin operator A_H is also symmetric and positive definite. The image space of an operator Q is denoted by $\mathcal{R}(Q)$. Using the theory of orthogonal projections a variational principle can be constructed for the coarse level correction operator $K_{h,H}$ [242].

Theorem C.6. *Let $A_h > 0$ (positive definiteness) and consider an arbitrary C/F-splitting and interpolation I_H^h of full rank. Then the coarse level correction operator $K_{h,H}$ is an orthogonal projector with regard to the energy product $(.,.)_1$. In particular:*

(i) $\mathcal{R}(K_{h,H}) \perp_1 \mathcal{R}(I_H^h)$ *or* $(A_h K_{h,H} u^h, I_H^h v^H)_E = 0$ *holds for all* u^h, v^H ,

(ii) $\|u^h + v^h\|_1^2 = \|u^h\|_1^2 + \|v^h\|_1^2$ *holds for* $u^h \in \mathcal{R}(K_{h,H})$ *and* $v^h \in \mathcal{R}(I_H^h)$,

(iii) $\|K_{h,H}\|_1 = 1$,

(iv) $\|K_{h,H} e^h\|_1 = \min_{e^H} \|e^h - I_H^h e^H\|_1$ *holds for all* e^h .

The fourth statement of the above proposition expresses a variational principle: Galerkin-based coarse-grid correction minimises the error in the energy norm with regard to all variants in $\mathcal{R}(I_H^h)$. As a consequence a two-level method never diverges as long as the smoother satisfies $\|S_h\|_1 \leq 1$. This is also valid for complete V-cycles with arbitrary C/F-hierarchies and arbitrary interpolation operators having full rank, as seen by recursive application of the following Lemma [242]:

Lemma C.7. *Replace the exact coarse-grid correction e^H in* (5.83) *by an arbitrary approximation $\tilde{e}^H$ with $\|e^H - \tilde{e}^H\|_1 \leq \|e^H\|_1$. Then the approximated coarse-grid correction operator $\tilde{K}_{h,H}$ satisfies the inequality $\|\tilde{K}_{h,H}\|_1 \leq 1$.*

This statement does not quantify the efficiency of a V-cycle but provides a minimum of robustness. In respect thereof, the truncation of interpolation operators as mentioned earlier is a "safe" process as long as it is performed *before* computing the Galerkin operator. In the worst case, convergence is downgraded but divergence is impossible. In contrast, the direct truncation of the Galerkin operator itself violates the variational principle and should only be used carefully.

(b) Post-smoothing and two-level convergence

Consider two-level convergence for symmetric positive definite matrices assuming $I_h^H = (I_H^h)^T$ so that the Galerkin operator remains symmetric and positive definite as well and that the coarse level correction operator fulfils the variational principle. Furthermore, assume that only one post-smoothing step is performed per cycle, so that a two-level operator K can be applied.

For the product SKe to become small, it is important that the smoothing operator S is able to reduce all vectors in $\mathcal{R}(K)$ efficiently. If the smoothing property is fulfilled, smoothing becomes more efficient if $\|e\|_2$ is small in comparison to $\|e\|_1$. A minimal requirement is that for all $e \in \mathcal{R}(K)$ the norm $\|e\|_2$ is bounded from below by $\|e\|_1$. This leads to the following proposition [242]:

Theorem C.8. *Let $A > 0$ and assume S to fulfil the smoothing property. Furthermore, the C/F-splitting and the interpolation are such that:*

$$\|Ke\|_1^2 \leq \tau \|Ke\|_2^2 \,, \tag{C.31}$$

with $\tau > 0$ independent of e. Then $\tau \geq \sigma$ and $\|SK\|_1 \leq \sqrt{1 - \sigma/\tau}$ hold.

In general the above coarsening and interpolation requirement, necessary to provide such a convergence estimate, is contained in the following lemma:

Lemma C.9. *If the C/F-splitting and interpolation I_{FC} are such that, for all e,*

$$\|e_F - I_{FC} e_C\|_{0,F}^2 \leq \tau \|e\|_1^2 \,, \tag{C.32}$$

is fulfilled with τ being independent of e, then (C.31) *is satisfied.*

In the above formula e_C and e_F denote restrictions of the vector e to either the set of C- or F-variables. Accordingly, I_{FC} denotes the interpolation matrix from the C to the F-variables and $\|\cdot\|_{0,F}$ relates to the 0-norm based on (C.20) except that the measured matrices and vectors are restricted to the F-variable set. Now, particularly for symmetric M-matrices, the convergence conditions can be formulated more practically:

Lemma C.10. *Let A be a symmetric M-matrix with $s_i = \sum_j a_{ij} \geq 0$. With fixed $\tau \geq 1$ choose a C/F-splitting such that for each $i \in F$ a set $P_i \subseteq C \cap N_i$ exists with:*

$$\sum_{k \in P_i} |a_{ik}| \geq \frac{1}{\tau} \sum_{j \in N_i} |a_{ij}| \,.$$

Then, the direct interpolation of Sec. C.4.3(a) fulfils condition (C.32) *and thus* (C.31).

In this way a direct relationship between convergence and interpolation quality is established: with an increasing coupling fraction $1/\tau$ of the interpolation variables P_i in proportion to the total coupling strength of the neighbours N_i, the interpolation quality of an F-variable i improves as well as the convergence. The convergence speed complies with the interpolation formula involving the weakest coupling fraction $1/\tau$.

The assumption of weak diagonal dominance in Lemma C.10 is sufficient but not necessary, a similar result can be stated for symmetric M-matrices the row sum of which is bounded from below. The above results can be generalised from direct to indirect interpolation.

D Application related calculations

This appendix is related to calculations required in Chapter 6, particularly including calculations for the spectral approach in Sec. 6.2, the analytic solution of film flow over a hemisphere, Sec. 6.3, and the turnstile-lobe effect, Sec. 6.4.

D.1 Film flow over periodically corrugated topography

In order to obtain the two systems of non-linear equations (6.40) and (6.41) in Sec. 6.2.3(b) various series representations and the evaluation of a number of inner products are required.

D.1.1 Calculations for the dynamic BC (6.26)

For the series representation of the dynamic boundary condition (6.26) the occurring g, Q and R terms, which are considered separately, have to be replaced by the expansions (6.35)-(6.39). For convenience the multiple sums are written down in a compact notation and, if not mentioned otherwise, all sum indices are assumed to run from $-\infty$ to ∞. Firstly, the expansion (6.35) gives the relationships:

$$\sigma\frac{\sqrt{1+g''^2}-1}{2} \approx \frac{\sigma}{2}\left(1+\frac{g''^2}{2}-1\right) = \frac{\sigma}{4}\sum_{k,l} l^2(k-l)^2 g_l g_{k-l} e^{\mathrm{i}kx}\,,$$

$$g'g'' = -\mathrm{i}\sum_{k,l} l^2(k-l) g_l g_{k-l} e^{\mathrm{i}kx}\,,$$

$$(g')^2 g'' = \sum_{k,l,n} l^2(k-l)(n-k) g_l g_{k-l} g_{n-k} e^{\mathrm{i}nx}\,,$$

$$g'^2 = -\sum_{k,l} l(k-l) g_l g_{k-l} e^{\mathrm{i}kx}\,,$$

$$gg'' = -\sum_{k,l} l^2 g_l g_{k-l} e^{\mathrm{i}kx}\,.$$

A little more involved but straightforward is the last term:

$$\mathrm{Im}(Q(1-\mathrm{i}g'')) = \frac{w_0}{2} - \frac{\mathrm{i}}{2}\sum_{k\neq 0}\left(e^{-kh}Q_k - e^{kh}\bar{Q}_{-k}\right)e^{\mathrm{i}kx} + \frac{u_0}{2}\sum_k k^2 g_k e^{\mathrm{i}kx}$$

$$+\frac{1}{2}\sum_{k\neq 0,l}(l-k)(l-2k)\left(e^{-kh}Q_k+e^{kh}\bar{Q}_{-k}\right)g_{l-k}e^{\mathrm{i}lx}$$
$$-\frac{i}{2}\sum_{k,l,n}k(l-k)(n-l)^2\left(e^{-kh}Q_k-e^{kh}\bar{Q}_{-k}\right)g_{l-k}g_{n-l}e^{\mathrm{i}nx}\,.$$

If all the above expansion formulas are inserted into (6.26), a system of equations can be constructed via $(e^{\mathrm{i}mx},(6.26))_0=0$ for $m\in\mathbb{Z}$ in which the inner product (6.31) is used. The required inner products are calculated separately:

$$\left(e^{\mathrm{i}mx},\frac{\sigma}{2}\left(\sqrt{1+g''^2}-1\right)\right)_0=\frac{\sigma}{4}\sum_l l^2(m-l)^2g_lg_{m-l}\,,$$
$$\left(e^{\mathrm{i}mx},g\right)_0=g_m\,,$$
$$\left(e^{\mathrm{i}mx},h^2g''\right)_0=-h^2m^2g_m\,,$$
$$\left(e^{\mathrm{i}mx},(g')^2g''/2\right)_0=\frac{1}{2}\sum_{k,l}l^2(k-l)(m-k)g_lg_{k-l}g_{m-k}\,,$$
$$\left(e^{\mathrm{i}mx},\cot(\alpha)g'^2/2\right)_0=-\frac{\cot\alpha}{2}\sum_l l(m-l)g_lg_{m-l}\,,$$
$$\left(e^{\mathrm{i}mx},\cot(\alpha)gg''\right)_0=-\cot\alpha\sum_l l^2g_lg_{m-l}\,,$$

and, again, the last term is more complicated:

$$\left(e^{\mathrm{i}mx},2\,\mathrm{Im}\left(Q\left(1-\mathrm{i}g''\right)\right)\right)_0$$
$$=-\mathrm{i}\left(e^{-mh}Q_m-e^{mh}\bar{Q}_{-m}\right)(1-\delta_{0m})+\sum_{k\neq 0}(m-k)(m-2k)\left(e^{-kh}Q_k+e^{kh}\bar{Q}_{-k}\right)g_{m-k}$$
$$-\mathrm{i}\sum_{k,l}k(l-k)(m-l)^2\left(e^{-kh}Q_k-e^{kh}\bar{Q}_{-k}\right)g_{l-k}g_{m-l}+w_0\delta_{0m}+u_0m^2g_m\,.$$

In summary a linear system of equations for the differential equation (6.26) is obtained by $(e^{\mathrm{i}mx},(6.26))_0=0$ resulting in (6.40).

D.1.2 Calculations for the dynamic BC (6.28)

For the series representation of the hybrid boundary condition (6.28) the same procedure as above is employed. Firstly, the following two additional relationships are required:

$$(g')^3=-\mathrm{i}\sum_{k,l,n}l(k-l)(n-k)g_lg_{k-l}g_{n-k}e^{\mathrm{i}nx}\,,$$
$$gg'=\mathrm{i}\sum_{k,l}lg_lg_{k-l}e^{\mathrm{i}kx}\,.$$

The occurring integral term in (6.28) can, by partial integration, be expanded as:

$$I := \int \left(gg'' + \frac{g'^2}{2} + \cot(\alpha) g \right) dx = \int \left(-\frac{1}{2} g'^2 + \cot(\alpha) g \right) dx + gg' \\ = \frac{\mathrm{i}}{2} \sum_{k \neq 0, l} \frac{l(l+k)}{k} g_l g_{k-l} e^{\mathrm{i}kx} - \mathrm{i} \cot \alpha \sum_{k \neq 0} \frac{g_k}{k} e^{\mathrm{i}kx} + C \,,$$

while the term $P + zQ$ gives:

$$\begin{aligned} P + zQ = {} & \frac{\psi_0 + hu_0 + \mathrm{i}hw_0}{2} + \sum_{k \neq 0} e^{-kh} \frac{(1+2hk)Q_k - R_k}{2k} e^{\mathrm{i}kx} + \frac{\mathrm{i}u_0 - w_0}{2} \sum_k k g_k e^{\mathrm{i}kx} \\ & + \mathrm{i} \sum_{k \neq 0, l} e^{-kh} (l-k) g_{l-k} \frac{(k-2hk)Q_k + R_k}{2} e^{\mathrm{i}lx} \\ & + \sum_{k,l,n} e^{-kh} l(l-k)(n-l) g_{l-k} g_{n-l} Q_k e^{\mathrm{i}nx} \,. \end{aligned}$$

Similar to the previous Sec. D.1.1 the following inner products are required for the discretisation of the hybrid boundary condition (6.28):

$$\begin{aligned} \left(e^{\mathrm{i}mx}, (h^2 - \mathrm{i}\sigma/2) g' \right)_0 &= \left(\mathrm{i}h^2 + \sigma/2 \right) k g_m \,, \\ \left(e^{\mathrm{i}mx}, \mathrm{i}I \right)_0 &= -\frac{1}{2} \sum_l \frac{l(l+m)}{m} g_l g_{m-l} \left(1 - \delta_{0m}\right) + \cot(\alpha) \frac{g_m}{m} \left(1 - \delta_{0m}\right) + \mathrm{i}C\delta_{0m} \,, \\ \left(e^{\mathrm{i}mx}, (\mathrm{i}\cot\alpha - 2) g'^3/6 \right)_0 &= \frac{\cot\alpha + 2\mathrm{i}}{6} \sum_{k,l} k(l-k)(m-l) g_k g_{l-k} g_{m-l} \,. \end{aligned}$$

Further calculations deliver:

$$\begin{aligned} \left(e^{\mathrm{i}mx}, 2\left(P + zQ\right) \right)_0 = {} & \left(\psi_0 + hu_0 + \mathrm{i}hw_0\right) \delta_{0m} + e^{-mh} \frac{(1+2hm)Q_m - R_m}{m} \left(1 - \delta_{0m}\right) \\ & + (\mathrm{i}u_0 - w_0) m g_m + \mathrm{i} \sum_{k \neq 0} e^{-kh} (m-k) g_{m-k} \big((k-2hk)Q_k + R_k \big) \\ & + 2 \sum_{k,l} e^{-kh} k(l-k)(m-l) g_{l-k} g_{m-l} Q_k \,, \end{aligned}$$

and:

$$\left(e^{\mathrm{i}mx}, \psi_s - \frac{2}{3} h^3 \right)_0 = \left(\psi_s - \frac{2}{3} h^3 \right) \delta_{0m} \,.$$

In summary a linear system of equations for the differential equation (6.28) is obtained by the inner product $(e^{\mathrm{i}mx}, (6.26))_0 = 0$ resulting in (6.41).

D.2 Film flow over non-planar surfaces

D.2.1 Boundary conditions for flow over a hemisphere

At the surface of the hemisphere, $r = r_0$, the no-penetration condition $\psi(r_0, \varphi) = 0$ and the no-slip condition $u_\varphi(r_0, \varphi) = 0$ imply:

$$\left[-\frac{r_0^3}{3}\Phi_2'(\vartheta) + r_0^2\left[\Phi_1''(\vartheta) + \Phi_1(\vartheta)\right] - r_0\Phi_3(\vartheta)\right]\sin\vartheta + \Phi_4(\vartheta) = 0\,, \tag{D.1}$$

$$r_0\Phi_2'(\vartheta) - 2\left[\Phi_1''(\vartheta) + \Phi_1(\vartheta)\right] + \frac{1}{r_0}\Phi_3(\vartheta) = 0\,, \tag{D.2}$$

allowing the two functions $\Phi_3(\vartheta)$ and $\Phi_4(\vartheta)$ to be expressed as combinations of $\Phi_1(\vartheta)$ and $\Phi_2(\vartheta)$ and their derivatives: taking the combination (6.53)$-$(D.1)$+r_0(r-r_0)\sin\vartheta$(D.2), the streamfunction simplifies to:

$$\eta\psi = (r - r_0)^2\left[\Phi_1''(\vartheta) + \Phi_1(\vartheta) - \frac{2r_0 + r}{3}\Phi_2'(\vartheta)\right]\sin\vartheta\,, \tag{D.3}$$

while the corresponding velocity profile reads:

$$\eta u_\vartheta = -2\left(1 - \frac{r_0}{r}\right)\left[\Phi_1''(\vartheta) + \Phi_1(\vartheta) - \frac{r + r_0}{2}\Phi_2'(\vartheta)\right]\,. \tag{D.4}$$

Next the kinematic boundary condition $\psi\left(r_0 f(\vartheta), \vartheta\right) = -\dot{V}_0$ implies:

$$\Phi_1'' + \Phi_1 - \frac{f+2}{3}r_0\Phi_2' = -\frac{\eta\dot{V}_0}{r_0^2\left(f-1\right)^2\sin\vartheta}\,. \tag{D.5}$$

Since the first integral of the dynamic boundary condition, (3.144), (3.145), refers to the gradient of the potential Φ, equations (6.50), (6.51) have to be considered. Due to the lubrication approximation the term $3\eta u_r/3$ in (6.51) is neglected again and the square root appearing in (3.144), (3.145) is approximated as $\sqrt{f(\vartheta)^2 + f'(\vartheta)^2} \approx f(\vartheta)$. These simplifications lead to the following two equations:

$$r_0 f(\vartheta)\Phi_2(\vartheta) - \Phi_1'(\vartheta) + \frac{\sigma}{2} - \frac{\varrho g r_0^2}{4}\left[f(\vartheta)^2\cos\vartheta + \int f(\vartheta)^2\mathrm{d}\vartheta\sin\vartheta\right] = 0, \tag{D.6}$$

$$\frac{f(\vartheta)^2-1}{2}r_0\Phi_2'(\vartheta) - \left[f(\vartheta)-1\right]\Phi_1''(\vartheta) + \Phi_1(\vartheta) - \frac{\sigma}{2}f'(\vartheta) - \frac{\varrho g r_0^2}{4}f(\vartheta)\cos\vartheta\int f(\vartheta)^2\mathrm{d}\vartheta = 0, \tag{D.7}$$

where (D.4) has been considered. Together with (D.5), they provide a set of three equations for three unknown functions $\Phi_1(\vartheta)$, $\Phi_2(\vartheta)$ and $f(\vartheta)$. A further reduction to a set of two equations is possible by eliminating $\Phi_1(\vartheta)$ as follows: subtracting (D.5) from (D.7) gives:

$$\frac{1}{6}\left(3f^2 + 2f + 1\right)r_0\Phi_2' - f\Phi_1'' - \frac{\sigma}{2}f' - \frac{\varrho g r_0^2}{4}f\cos\vartheta\int f^2\mathrm{d}\vartheta = \frac{\eta\dot{V}_0}{r_0^2\left(f-1\right)^2\sin\vartheta}\,, \tag{D.8}$$

while, taking the derivative of (D.6), Φ_1'' can be expressed as:

$$\Phi_1'' = r_0 f \Phi_2' + r_0 f' \Phi_2 - \frac{\varrho g r_0^2}{4} \left[2ff' + \int f^2 \mathrm{d}\vartheta \right] \cos\vartheta \,. \tag{D.9}$$

Eliminated Φ_1'' in (D.8) according to (D.9), yields:

$$\frac{1}{6} \underbrace{\left(-3f^2 + 2f + 1 \right)}_{-(3f+1)(f-1)} r_0 \Phi_2' - r_0 f f' \Phi_2 - \frac{\sigma}{2} f' + \frac{\varrho g r_0^2}{2} f^2 f' \cos\vartheta = \frac{\eta \dot{V}_0}{r_0^2 \left(f-1\right)^2 \sin\vartheta} \,, \tag{D.10}$$

as a pure ODE containing only Φ_2 and f. Inserting (D.9) into (D.5), delivers:

$$\Phi_1 + \frac{2}{3}(f-1) r_0 \Phi_2' + r_0 f' \Phi_2 - \frac{\varrho g r_0^2}{4} \left[2ff' + \int f^2 \mathrm{d}\vartheta \right] \cos\vartheta = -\frac{\eta \dot{V}_0}{r_0^2 \left(f-1\right)^2 \sin\vartheta} \,, \tag{D.11}$$

Taking the derivative of (D.11) and adding (D.6) yields:

$$\begin{aligned} &\frac{2}{3}(f-1) r_0 \Phi_2'' + \frac{5}{3} f' r_0 \Phi_2' + (f'' + f) r_0 \Phi_2 + \frac{\sigma}{2} \\ &\quad + \frac{\varrho g r_0^2}{2} \left[f f' \sin\vartheta - \left(f'^2 + f f'' + f^2 \right) \cos\vartheta \right] = \eta \dot{V}_0 \frac{2f' + (f-1)\cot\vartheta}{r_0^2 \left(f-1\right)^3 \sin\vartheta} \,, \end{aligned} \tag{D.12}$$

as a second ODE, to accompany (D.10), containing only Φ_2 and f. It is convenient to add the derivative of (D.10) to (D.12), leading to:

$$\begin{aligned} &-\frac{1}{2} \left(f-1\right)^2 r_0 \Phi_2'' - 2 r_0 (f-1) f' \Phi_2' - \left[f'^2 - f + (f-1) f'' \right] r_0 \Phi_2 + \frac{\sigma}{2}(1 - f'') \\ &\qquad + \frac{\varrho g r_0^2}{2} \left(\left[(2f-1) f'^2 + (f-1) f f'' - f^2 \right] \cos\vartheta - (f-1) f f' \sin\vartheta \right) = 0 \,. \end{aligned} \tag{D.13}$$

Via (D.10), (D.13) the set of equations is reduced to two equations for two unknown functions.

D.2.2 Comparison with heuristic considerations

The paper [246] utilises a form of the mass conservation obtained from the local mass conservation by integration with respect to the coordinate perpendicular to the flow direction, which the authors call *local mass conservation*. Using spherical coordinates for the flow over a hemisphere, the mean flow direction is the ϑ-direction, while the radial direction is perpendicular to the flow. Thus one obtains:

$$\partial_t h + \frac{1}{r \sin\vartheta} \partial_\vartheta \left(\sin\vartheta q \right) = 0 \,,$$

where q is assumed to be the flow rate of the Nusselt film flow at a given polar angle ϑ, i.e.:

$$q = \frac{\varrho g \sin\vartheta h^3}{3\eta} \,, \tag{D.14}$$

leading to the evolution equation:

$$\partial_t h + \frac{\varrho g}{3\eta r_0 \sin\vartheta}\partial_\vartheta \left(h^3 \sin^2\vartheta\right) = 0\,,$$

for the film thickness $h(\vartheta, t)$ in the unsteady case. For a steady flow, $\partial_t h = 0$, it follows from the above equation that:

$$h^3 \sin^2\vartheta = \text{const}\,,$$

in full accordance with the asymptotic result (6.65).

D.3 Turnstile lobe effect

D.3.1 General considerations for small Reynolds numbers

Note that in two dimensions the vorticity ω is given by $2\omega = \Delta\Psi$, and the corresponding evolution equation for ω, the vortex transport equation, is obtained from the imaginary part of (3.13), by applying the Laplacian $\Delta = 4\partial^2/\partial\xi\partial\bar{\xi}$, as:

$$\mathrm{Re}\frac{\partial\omega}{\partial t} - \Delta\omega = -\frac{1}{2}\Delta^2 \mathrm{Im}\chi\ ; \tag{D.15}$$

while taking the derivative of (3.12) with respect to ξ twice, leads to:

$$\Delta^2\chi = -\frac{\mathrm{Re}}{2}\frac{\partial^2 u^2}{\partial\xi^2}\,. \tag{D.16}$$

The Reynolds number appears in both of the above equations but with a different physical meaning: in equation (D.16) inertia is present by accounting for quadratic terms, whereas in equation (D.15) it is related to the unsteady character of the flow.

By assuming that inertia can be neglected due to physical reasoning while retaining the unsteady character of the flow, leads to the following simplified vortex transport equation:

$$\mathrm{Re}\frac{\partial\omega}{\partial t} - \Delta\omega = 0\,. \tag{D.17}$$

By taking the time derivative of (D.17), multiplying the result with Re and utilising (D.17) again, gives:

$$0 = \mathrm{Re}^2\frac{\partial^2\omega}{\partial t^2} - \Delta\left[\mathrm{Re}\frac{\partial\omega}{\partial t}\right] = \mathrm{Re}^2\frac{\partial^2\omega}{\partial t^2} - \Delta^2\omega\,. \tag{D.18}$$

Finally, noting that $\Delta^2\omega = \mathcal{O}\left(\mathrm{Re}^2\right)$ for small Reynolds numbers, and remembering that $2\omega = \Delta\Psi$, leads to:

$$\Delta^3\Psi = \mathcal{O}\left(\mathrm{Re}^2\right)\,.$$

D.3.2 Fourier discretization

(a) Base flow: steady Stokes flow ($h_1 = 0$)

By inserting the series representation (6.85), (6.86) into condition (6.82) at the upper corrugated surface, it follows that:

$$2h_0 B\delta_{0n} + \exp(-nh_0)\left[(1-2nh_0)\,Q_n - 2nR_n\right] + \exp(-nh_0)\bar{Q}_{-n} = \delta_{0n}\,. \tag{D.19}$$

For $n = 0$ the constant B can be expressed in terms of Q_0 via:

$$B = \frac{1}{2h_0}\left[1 - 2\mathrm{Re}Q_0\right]\,; \tag{D.20}$$

while for $n > 0$, the negative-indexed Q-coefficients can be expressed in terms of their positive-indexed counterparts.:

$$\bar{Q}_{-n} = \exp(-2nh_0)\left[2nR_n - (1-2nh_0)\,Q_n\right]\,. \tag{D.21}$$

For the case $n < 0$ use is made of the complex conjugate of the equations and the substitution $n \to -n$, leading finally to:

$$\bar{R}_{-n} = -\exp(-2nh_0)\left[(1+2nh_0)\,R_n + 2nh_0^2 Q_n\right]\,, \tag{D.22}$$

enabling, as above, the negative-indexed R-coefficients to be expressed in terms of their positive-indexed counterparts.

Next, the series representation (6.85), (6.86) is applied to boundary condition (6.80) related to the lower corrugated surface and considering equations (D.20), (D.21) and (D.22). Furthermore the gauging condition $\mathrm{Im}Q_0 = 0$ is added to the set of equations and the Fourier decomposition [1]:

$$\exp\left(ka\cos x\right) = \sum_{n=-\infty}^{+\infty} \mathrm{I}_n(ka)\exp(\mathrm{i}nx)\,, \tag{D.23}$$

is used with I_n being the modified Bessel functions of order n. Finally, a set of linear algebraic equations follow:

$$\sum_{k=0}^{\infty}\left\{4k^2\frac{\mathrm{I}^{(1)}_{n-k} - h_0\mathrm{I}^{(0)}_{n-k}}{\exp(2kh_0)}\bar{R}_k - 2k\left[\mathrm{I}^{(0)}_{nk} - \frac{\mathrm{I}^{(0)}_{-n-k}}{\exp(2kh_0)}\right]R_k + \left[\mathrm{I}^{(0)}_{nk} - 2k\mathrm{I}^{(1)}_{nk} - \frac{1-2kh_0}{\exp(2kh_0)}\mathrm{I}^{(0)}_{-n-k}\right]Q_k\right.$$
$$\left. + \left[\mathrm{I}^{(0)}_{-nk} - \frac{1-2kh_0+4k^2h_0^2}{\exp(2kh_0)}\mathrm{I}^{(0)}_{n-k} + \frac{2k(1-2kh_0)}{\exp(2kh_0)}\mathrm{I}^{(0)}_{n-k}\right]\bar{Q}_k\right\} = \frac{a}{2h_0}\left[\delta_{1n} + \delta_{-1n}\right]\,, \tag{D.24}$$

where $\mathrm{I}_{nk}^{(0)} := \mathrm{I}_{n-k}(ka)$ and $\mathrm{I}_{nk}^{(p+1)} := -a\left[\mathrm{I}_{n-1\,k}^{(p)} + \mathrm{I}_{n+1\,k}^{(p)}\right]/2$ are used as abbreviations $(p = 0, 1)$. Note that in (D.24) the index n goes from $-N$ to N.

(b) Perturbation

By inserting (6.84)-(6.84) in (6.79), the complex conjugate of the velocity reads:

$$\begin{aligned}\bar{u}_\mathrm{u} = &\left\{2\mathrm{i}\left[r^{+\prime} + yq^{+\prime} + \mathrm{Re}\,\frac{y^2}{2}p^{+\prime}\right] + q^+ + \overline{q^-} + \mathrm{Re}\,y\left[p^+ + \overline{p^-}\right]\right\}\exp(+\mathrm{i}t)\\ &+\left\{2\mathrm{i}\left[r^{-\prime} + yq^{-\prime} + \mathrm{Re}\,\frac{y^2}{2}p^{-\prime}\right] + q^- + \overline{q^+} + \mathrm{Re}\,y\left[p^- + \overline{p^+}\right]\right\}\exp(-\mathrm{i}t)\,.\end{aligned}$$

Next, by applying the series representation (6.85), (6.86) and (6.87), the following linear algebraic set of equations is obtained $(n \neq 0)$ from the boundary condition (6.81):

$$\begin{aligned}\sum_{\substack{k=-N\\k\neq 0}}^{N}&\left\{\left[\mathrm{I}_{nk}^{(0)} - 2k\mathrm{I}_{nk}^{(1)} \pm \frac{\mathrm{iRe}}{2k}\left(k\mathrm{I}_{nk}^{(2)} - \mathrm{I}_{nk}^{(1)}\right)\right]q_k^\pm + \left[\mathrm{I}_{-nk}^{(0)} \mp \frac{\mathrm{iRe}}{2k}\mathrm{I}_{-nk}^{(1)}\right]\bar{q}_k^\mp - 2k\mathrm{I}_{nk}^{(0)}r_k^\pm\right\}\\ &- a\mathrm{Re}\left(\delta_{1n} + \delta_{-1n}\right)p_0 = 0\,,\end{aligned} \tag{D.25}$$

for the coefficients $q_k^\pm, r_k^\pm$ and p_0. For convenience the derivative of (6.82) with respect to x is taken, implying the identity $\partial\bar{u}_\mathrm{s}/\partial\xi = -\partial\bar{u}_\mathrm{s}/\partial\bar{\xi}$ at $\xi = x + \mathrm{i}h_0$, in order to simplify the boundary condition (6.83) together with (6.78) as follows:

$$\bar{u}_\mathrm{u}\left(x + \mathrm{i}h_0, x - \mathrm{i}h_0, t\right) = -2\mathrm{i}\frac{\partial\bar{u}_\mathrm{s}}{\partial\bar{\xi}}\cos(x-t) = 2\left[B - 2\mathrm{Im}Q_\mathrm{s}'\right]\cos(x-t)\,. \tag{D.26}$$

Applying the series representation (6.85), (6.86) and (6.87) again, the boundary condition (6.83) implies the following linear algebraic set of equations $(n \neq 0)$:

$$\frac{\left[1 - 2nh_0 \pm \frac{\mathrm{i}h_0\mathrm{Re}}{2n}\left(nh_0 - 1\right)\right]q_n^\pm - 2nr_n^\pm}{\exp(nh_0)} + \exp(nh_0)\left[1 \pm \frac{\mathrm{i}h_0\mathrm{Re}}{2n}\right]\bar{q}_{-n}^\mp = b_n^\pm\,, \tag{D.27}$$

$$2h_0\,\mathrm{Re}\,p_0 = b_0^+\,, \tag{D.28}$$

for the coefficients $q_k^\pm, r_k^\pm$ and p_0, where the inhomogeneity $b_n^\pm$ is calculated according to:

$$b_n^\pm := B\delta_{\mp 1\,n} - \frac{1}{\pi}\int_{-\pi}^{+\pi}\mathrm{Im}Q_\mathrm{s}'\left(x + \mathrm{i}h_0\right)\exp\left(-\mathrm{i}[n \pm 1]x\right)\mathrm{d}x\,,$$

from the coefficient B and the function Q_s of the base solution.

Bibliography

[1] Abramowitz M., and Stegun I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, volume 55 of *Applied Mathematics Series*. US National Bureau of Standards, 1970.

[2] Agmon S., Douglis A., and Nirenberg L. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I. *Com. on Pure and Appl. Math.*, **12**(4):623–727, 1959.

[3] Agmon S., Douglis A., and Nirenberg L. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II. *Com. on Pure and Appl. Math.*, **17**(1):35–92, 1964.

[4] Agranovich M. S., and Vishik M. I. Elliptic problems with a parameter and parabolic problems of general type. *Russian Mathematical Surveys*, **19**(3):53–157, 1964.

[5] Aharonov Y., and Bohm D. Significance of electromagnetic potentials in the quantum theory. *Physical Review*, **115**(3):485, 1959.

[6] Aharonov Y., Cohen E., and Rohrlich D. Nonlocality of the Aharonov-Bohm effect. *Physical Review A*, **93**(4):042110, 2016.

[7] Aksel N., and Schörner M. Films over topography: from creeping flow to linear stability, theory, and experiments, a review. *Acta Mechanica*, **229**(4):1453–1482, Apr 2018. ISSN 1619-6937.

[8] Antanovskii L. K. Boundary integral equations in quasi-steady problems of capillary fluid mechanics. Part 2: Application of the stress-stream function. *Meccanica - J. Ital. Assoc. Theoret. Appl. Mech.*, **26(1)**:59–65, 1991.

[9] Anthony K. H. Unification of continuum mechanics and thermodynamic by means of Lagrange-formalism – present status of the theory and presumable applications. *Archives of Mechanics*, **41**(4):511–534, 1989.

[10] Anthony K. H. Phenomenological thermodynamics of irreversible processes within Lagrange formalism. *Acta Physica Hungarica*, **67**(3-4):321–340, 1990.

[11] Anthony K. H. Hamilton's action principle and thermodynamics of irreversible processes – a unifying procedure for reversible and irreversible processes. *J. Non-Newt. Fluid Mech.*, **96**:291–339, 2001.

[12] Apel T., Sändig A. M., and Whiteman J. R. Graded mesh refinement and error estimates for finite element solutions of elliptic boundary value problems in non-smooth domains. *Math. Meth. in the Appl. Sciences*, **19**(1):63–85, 1996.

[13] Arnaudon M., and Cruzeiro A. B. Lagrangian Navier-Stokes diffusions on manifolds: variational principle and stability. *Bulletin des Sciences Mathématiques*, **136** (8):857–881, 2012.

[14] Arnaudon M., and Cruzeiro A. B. Stochastic Lagrangian flows and the Navier-Stokes equations. In *Stochastic Analysis: A Series of Lectures*, pages 55–75. Springer, 2015.

[15] Arnaudon M., Cruzeiro A. B., and Galamba N. Lagrangian Navier-Stokes flows: a stochastic model. *J. Phys. A*, **44**(17):175501, 2011.

[16] Ash R. L., and Zardadkhan I. R. Non-equilibrium pressure control of the height of a large-scale, ground-coupled, rotating fluid column. *Phys. Fluids*, **25**(5):053101, 2013.

[17] Ash R. L., Zardadkhan I., and Zuckerwar A. J. The influence of pressure relaxation on the structure of an axial vortex. *Phys. Fluids*, **23**(7):073101, 2011.

[18] Babuška I., and Rosenzweig M. B. A finite element scheme for domains with corners. *Numerische Mathematik*, **20**(1):1–21, 1972.

[19] Babuška I., Osborn J., and Pitkäranta J. Analysis of mixed methods using mesh dependent norms. *Mathematics of Computation*, **35**(152):1039–1062, 1980.

[20] Balkovsky E. Some notes on the Clebsch representation for incompressible fluids. *Physics Letters A*, **186**(1):135–136, 1994.

[21] Bartels S., Carstensen C., and Hecht A. P2Q2Iso2D = 2D Isoparametric FEM in Matlab. *J. Comput. Appl. Math.*, **192**(2):219–250, 2006.

[22] Batchelor G. K. *An Introduction to Fluid Dynamics*. Cambridge Mathematical Library. Cambridge University Press, 2000.

[23] Bateman H. Notes on a differential equation which occurs in the two-dimensional motion of a compressible fluid and the associated variational problems. *Proc. R. Soc. Lond. A*, **125**(799):598–618, 1929.

[24] Bateman H. On dissipative systems and related variational principles. *Physical Review*, **38**(4):815, 1931.

[25] Bathe K. J. *Finite-Elemente-Methoden*, volume 2. Springer Berlin, 2002.

[26] Bauer P. S. Dissipative dynamical systems i. *Proceedings of the National Academy of Sciences*, **17**(5):311–314, 1931.

[27] Bendali A., Dominguez J. M., and Gallic S. A variational approach for the vector potential formulation of the Stokes and Navier-Stokes problems in three dimensional domains. *J. Math. Analysis and Appl.*, **107**(2):537–560, 1985. ISSN 0022-247X.

[28] Benzi M., Golub G. H., and Liesen J. Numerical solution of saddle point problems. *Acta Numerica*, **14**:1–137, 2005.

[29] Bernardi C. Optimal finite-element interpolation on curved domains. *SIAM J. Num. Analysis*, **26**(5):1212–1240, 1989.

[30] Bernardi C., Girault V., and Maday Y. Mixed spectral element approximation of the navier-stokes equations in the stream-function and vorticity formulation. *IMA J. Num. Analysis*, **12**(4):565–608, 1992.

[31] Bertrand F., Münzenmaier S., and Starke G. First-order system least squares on curved boundaries: Higher-order Raviart–Thomas elements. *SIAM J. Num. Analysis*, **52**(6):3165–3180, 2014.

[32] Blum H. On the approximation of linear elliptic systems on polygonal domains. In *Singularities and Constructive Methods for Their Treatment*, pages 28–37. Springer, 1985.

[33] Blum H., and Dobrowolski M. On finite element methods for elliptic equations on domains with corners. *Computing*, **28**(1):53–63, 1982.

[34] Blum H., Rannacher R., and Leis R. On the boundary value problem of the biharmonic operator on domains with angular corners. *Math. Methods in the Appl. Sci.*, **2**(4):556–581, 1980.

[35] Bochev P., Cai Z., Manteuffel T., and McCormick S. Analysis of velocity-flux first-order system least-squares principles for the Navier–Stokes equations: Part I. *SIAM J. Num. Analysis*, **35**(3):990–1009, 1998.

[36] Bochev P. B., and Gunzburger M. D. Analysis of least squares finite element methods for the Stokes equations. *Math. of Comput.*, **63**(208):479–506, 1994.

[37] Bochev P. B., and Gunzburger M. D. *Least-Squares Finite Element Methods*, volume 166 of *Applied Mathematical Sciences*. Springer, New York, 2009.

[38] Bolton P., and Thatcher R. W. A least-squares finite element method for the Navier-Stokes equations. *J. Comput. Phys.*, **213**:174–183, 2006.

[39] Bontozoglou V. Laminar film flow along a periodic wall. *CMES (Computer Modelling in Engineering & Sciences)*, **1**(2):133–142, 2000.

[40] Bontozoglou V., and Papapolymerou G. Laminar film flow down a wavy incline. *Int. J. Multiphase Flow*, **23**(1):69–79, 1997.

[41] Boozer A. H. Magnetic field line Hamiltonian. Technical report, Princeton Univ., 1985.

[42] Boyer T. H. Does the Aharonov–Bohm effect exist? *Foundations of Physics*, **30** (6):893–905, 2000.

[43] Braess D. *Multigrid Methods II*, volume 1228 of *Lecture Notes in Mathematics*, chapter On the Combination of the Multigrid Method and Conjugate Gradients, pages 52 – 64. Springer Verlag, 1986.

[44] Braess D. *Finite Elemente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer-Verlag, 2013.

[45] Brandt A. Algebraic multigrid theory: The symmetric case. *Appl. Math. Comp.*, **4**:23–56, 1986.

[46] Brandt A., McCormick S., and Ruge J. Algebraic multigrid (AMG) for sparse matrix equations. In Evans D., editor, *Sparsity and Its Applications*. Cambridge University Press, 1984.

[47] Brenner H. Navier–Stokes revisited. *Physica A: Stat. Mech. and Appl.*, **349**(1-2): 60–132, 2005.

[48] Brenner S., and Scott R. *The Mathematical Theory of Finite Element Methods*, volume 15. Springer Science & Business Media, 2007.

[49] Bretherton F. P. A note on Hamilton's principle for perfect fluids. *J. Fluid Mech.*, **44**(1):19–31, 1970.

[50] Brezina M., Cleary A. J., Falgout R. D., Henson V. E., Jones J. E., Manteuffel T. A., McCormick S. F., and Ruge J. W. Algebraic multigrid based on element interpolation (AMGe). *SIAM J. Sci. Comput.*, **22(5)**:1570–1592, 2000.

[51] Brezina M., Falgout R. D., MacLachlan S., Manteuffel T. A., McCormick S. F., and Ruge J. W. Adaptive smoothed aggregation (αSA) multigrid. *SIAM Review: SIGEST*, **47(2)**:317–346, 2005.

[52] Brezina M., Falgout R. D., MacLachlan S., Manteuffel T. A., McCormick S. F., and Ruge J. W. Adaptive algebraic multigrid. *SIAM J. Sci. Comput.*, **27(4)**:1261–1286, 2006.

[53] Brezis H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer Science & Business Media, 2010.

[54] Brezzi F., Rappaz J., and Raviart P. A. Finite dimensional approximation of nonlinear problems. *Numerische Mathematik*, **36**(1):1–25, 1980.

[55] Brezzi F., Boffi D., Demkowicz L., Durán R., Falk R., and Fortin M. *Mixed Finite Elements, Compatibility Conditions, and Applications*. Springer, 2008.

[56] Briggs W. L., Henson V. E., and McCormick S. F. *A Multigrid Tutorial*. SIAM (Society for Industrial and Applied Mathematics), Philadelphia, 2nd edition, 2000.

[57] Cai Z., Kim S., and Woo G. A finite element method using singular functions for the poisson equation: crack singularities. *Num. Lin. Algebra with Appl.*, **9**(6-7): 445–455, 2002.

[58] Calkin M. G. An action principle for magnetohydrodynamics. *Canadian J. Phys.*, **41**(12):2241–2251, 1963.

[59] Cassidy M. *A Spectral Method for Viscoelastic Extrudate Swell.* PhD thesis, Univ. of Wales, Aberystwyth, 1996.

[60] Chan R. H., DeLillo T. K., and Horn M. A. The numerical solution of the biharmonic equation by conformal mapping. *SIAM J. Sci. Comput.*, **18**(6):1571–1582, 1997.

[61] Chandrasekhar S. *Hydrodynamic and Hydromagnetic Stability.* Courier Corporation, 2013.

[62] Ciarlet P. G., and Raviart P. A. A mixed finite element method for the biharmonic equation. In *Proceedings of Symposium on Mathematical Aspects of Finite Elements in PDE*, pages 125–145, 1974.

[63] Cipriano F., and Cruzeiro A. B. Navier-Stokes equation and diffusions on the group of homeomorphisms of the torus. *Com. Math. Phys.*, **275**(1):255–269, 2007. ISSN 1432-0916.

[64] Clebsch A. Ueber die Integration der hydrodynamischen Gleichungen. *J. f. d. reine u. angew. Math.*, **56**:1–10, 1859.

[65] Clees T. *AMG Strategies for PDE Systems with Applications in Industrial Semiconductor Simulation.* PhD thesis, Universität zu Köln, 2005.

[66] Clees T., and Ganzer L. An efficient algebraic multi-grid solver strategy for adaptive implicit methods in oil reservoirs simulation. In *Proc. 2007 SPE Reservoir Simulation Symposium*, number SPE 105789, Houston, Februar 2007.

[67] Clees T., Samrowski T., Zimmermann M. L., and Weigel R. An automatic multi-level solver switching strategy for peec-based emc simulation. In *Electromagnetic Compatibility (EMC Zurich 2007). 18th International Zurich Symposium*, München, September 2007.

[68] Coleman C. J. A finite element routine for solving non-Newtonian flows. *J. Non-Newtonian Fluid Mech.*, **7**(4):289–301, 1980.

[69] Coleman C. J. A contour integral formulation of plane creeping Newtonian flow. *Quart. J. Mech. Appl. Math.*, **34**:453–464, 1981.

[70] Coleman C. J. On the use of complex variables in the analysis of flows of an elastic fluid. *J. Non-Newtonian Fluid Mech.*, **15**(2):227–238, 1984.

[71] Constantin P., Iyer G., et al. A stochastic-Lagrangian approach to the Navier–Stokes equations in domains with boundary. *The Annals of Applied Probability*, **21** (4):1466–1492, 2011.

[72] Costabel M., and Dauge M. Stable asymptotics for elliptic systems on plane domains with corners: Stable asymptotics for elliptic systems. *Com. in Part. Different. Equat.*, **19**(9-10):1677–1726, 1994.

[73] Costabel M., Dauge M., and Nicaise S. Analytic regularity for linear elliptic systems in polygons and polyhedra. *Math. Models and Meth. in Appl. Sci.*, **22**(08): 1250015, 2012.

[74] Cotter C. J., and Holm D. D. Continuous and discrete Clebsch variational principles. *Found. Comput. Math.*, **9**:221–242, 2009.

[75] Cox C. L., and Fix G. J. On the accuracy of least squares methods in the presence of corner singularities. *Computers & Mathematics with Applications*, **10**(6):463–475, 1984.

[76] Cresson J. Fractional embedding of differential operators and Lagrangian systems. *J. Math. Phys.*, **48**(3):033504, 2007.

[77] Darrigol O. *Worlds of Flow: A History of Hydrodynamics from the Bernoullis to Prandtl*. Oxford University Press, 2005.

[78] Dauth M., Schörner M., and Aksel N. What makes the free surface waves over topographies convex or concave? a study with Fourier analysis and particle tracking. *Phys. Fluids*, **29**(9):092108, 2017.

[79] Davey K., and Darvizeh R. Neglected transport equations: extended Rankine-Hugoniot conditions and J-integrals for fracture. *Continuum Mechanics and Thermodynamics*, **28**(5):1525–1552, 2016. ISSN 1432-0959.

[80] Davis D. R. The inverse problem of the calculus of variations in higher space. *Transact. American Math. Soc.*, **30**(4):710–736, 1928.

[81] De Coster C., Nicaise S., and Sweers G. Solving the biharmonic Dirichlet problem on domains with corners. *Mathematische Nachrichten*, **288**(8-9):854–871, 2015.

[82] de Lagrange J. L. Mémoire sur la théorie du mouvement des fluides. *Oeuvres complètes*, **695**, 1781.

[83] Dean B., and Bhushan B. Shark-skin surfaces for fluid-drag reduction in turbulent flow: a review. *Phil. Transact. Roy. Soc. Lond. A*, **368**(1929):4775–4806, 2010.

[84] Didenko V. D., and Helsing J. Stability of the Nyström method for the Sherman–Lauricella equation. *SIAM J. Num. Analysis*, **49**(3):1127–1148, 2011.

[85] Dijkstra H. A., Wubs F. W., Cliffe A. K., Doedel E., Dragomirescu I. F., Eckhardt B., Gelfgat A. Y., Hazel A. L., Lucarini V., Salinger A. G., et al. Numerical bifurcation methods and their application to fluid dynamics: analysis beyond simulation. *Com. Comput. Phys.*, **15**(1):1–45, 2014.

[86] Ding H., Shu C., Yeo K. S., and Xu D. Numerical computation of three-dimensional incompressible viscous flows in the primitive variable form by local multiquadric differential quadrature method. *Comput. Methods Appl. Mech. Engrg.*, **195**:516–533, 2006.

[87] Dreisigmeyer D. W., and Young P. M. Nonconservative Lagrangian mechanics: a generalized function approach. *J. Phys. A*, **36**(30):8297, 2003.

[88] Eckart C. Variation principles of hydrodynamics. *Phys. Fluids*, **3**(3):421–427, 1960.

[89] Elman H. C., Silvester D. J., and Wathen A. J. *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*. Numerical Mathematics & Scientific Computation, 2014.

[90] Erturk E., Corke T. C., and Gökçöl C. Numerical solutions of 2d steady incompressible driven cavity flow at high Reynolds numbers. *Int. J. Num. Meth. Fluids*, **48**(7):747–774, 2005.

[91] Esquivelzeta-Rabell F. M., Figueroa-Espinoza B., Legendre D., and Salles P. A note on the onset of recirculation in a 2D Couette flow over a wavy bottom. *Phys. Fluids*, **27**(1), 2015. ISSN 1070-6631.

[92] Evans L. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010. ISBN 9780821849743.

[93] Falgout R. D., and Yang U. M. hypre: A library of high performance preconditioners. In *Lecture Notes in Computer Science*, volume 2331, pages 632–641. Springer Verlag, 2002.

[94] Feireisl E., and Vasseur A. New perspectives in fluid dynamics: Mathematical analysis of a model proposed by Howard Brenner. In *New Directions in Mathematical Fluid Mechanics*, pages 153–179. Springer, 2009.

[95] Ferziger J. H., Peric M., and Leonard A. *Computational Methods for Fluid Dynamics*. Springer, Berlin Heidelberg, 3 edition, 2002.

[96] Feynman R., Leighton R., and Sands M. *The Feynman Lectures on Physics*, volume 2. Addison-Wesley, 1963. ISBN 9780201021165.

[97] Finlayson B., and Scriven L. The method of weighted residuals – a review. *Appl. Mech. Rev*, **19**(9):735–748, 1966.

[98] Finlayson B. A. Existence of variational principles for the Navier-Stokes equation. *Phys. Fluids*, **15**(6):963–967, 1972.

[99] Finlayson C. *The Method of Weighted Residuals and Variational Principles, with Application in Fluid Mechanics, Heat and Mass Transfer.* Mathematics in Science and Engineering. Elsevier Science, 1972. ISBN 9780080955964.

[100] Fries C., and Manhartsgruber B. A moving piston boundary condition including gap flow in OpenFOAM. *WSEAS Trans. Fluid Mech.*, **10**:95–104, 2015.

[101] Frisch U. *Turbulence: The Legacy of A. N. Kolmogorov.* Cambridge University Press, 1995. ISBN 9781139935975.

[102] Fukagawa H., and Fujitani Y. A variational principle for dissipative fluid dynamics. *Progress of Theoretical Physics*, **127**(5):921–935, 2012.

[103] Gaskell P., Summers J., Thompson H., and Savage M. Creeping flow analyses of free surface cavity flows. *Theoret. Comput. Fluid Dynamics*, **8**(6):415–433, 1996.

[104] Gaskell P., Thompson H., and Savage M. A finite element analysis of steady viscous flow in triangular cavities. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, **213**(3):263–276, 1999.

[105] Gaskell P. H., Jimack P. K., Sellier M., Thompson H. M., and Wilson M. C. T. Gravity-driven flow of continuous thin liquid films on non-porous substrates with topography. *J. Fluid Mech.*, **509**:253–280, 2004.

[106] Gazzola F., Grunau H., and Sweers G. *Polyharmonic Boundary Value Problems: Positivity Preserving and Nonlinear Higher Order Elliptic Equations in Bounded Domains.* Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2010.

[107] Girault V., and Raviart P. A. *Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms*, volume 5. Springer Science & Business Media, 2012.

[108] Goldstein H., Poole C. P., and Safko J. L. *Klassische Mechanik.* WILEY-VCH, 2006. ISBN 978-3-527-40589-3.

[109] Gradshteyn I. S., and Ryzhik I. M. *Table of Integrals, Series, and Products.* Elsevier/Academic Press, Amsterdam, 7. edition, 2007.

[110] Graves R. W. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences. *Bull. Seism. Soc. Am.*, **86**:1091–1106, 1996.

[111] Greengard L., Kropinski M. C., and Mayo A. Integral equation methods for Stokes flow and isotropic elasticity in the plane. *J. Comput. Phys.*, **125**:403–404, 1996.

[112] Greve R., and Blatter H. *Dynamics of Ice Sheets and Glaciers.* Springer Science & Business Media, 2009.

[113] Grmela M. Bracket formulation of dissipative fluid mechanics equations. *Physics Letters A*, **102**:355–358, 1984.

[114] Guo B., and Babuska I. On the regularity of elasticity problems with piecewise analytic data. *Advances in Appl. Math.*, **14**(3):307–347, 1993.

[115] Guo B., and Schwab C. Analytic regularity of Stokes flow on polygonal domains in countably weighted Sobolev spaces. *J. Comput. Appl. Math.*, **190**(1-2):487–519, 2006.

[116] Gurtin M. Variational principles for linear initial-value problems. *Quarterly J. Appl. Math.*, **22**(3):252–256, 1964.

[117] Haas A. *Influence of topography on flow structure and temperature distribution in viscous flows*. PhD thesis, University of Bayreuth, 2010.

[118] Hackbusch W. *Theorie und Numerik elliptischer Differentialgleichungen*. Springer, 1986.

[119] Hackbusch W. *Multigrid Methods and Applications*, volume 4. Springer Science & Business Media, 2013.

[120] Hante F. M., Leugering G., and Seidman T. I. Modeling and analysis of modal switching in networked transport systems. *Appl. Math. and Optimization*, **59**(2): 275–292, 2009. ISSN 1432-0606.

[121] Heaviside O. Electrical papers (2 volumes, collected works),. The Electrician Printing and Publishing Co., 1892.

[122] Heaviside O. *Electromagnetic Theory*, volume 1. The Electrician Printing and Publishing Co., 1894.

[123] Helmholtz H. Über Integrale der hydrodynamischen Gleichungen welche den Wirbelbewegungen entsprechen. *Crelles Journal*, **55**:25–55, 1858.

[124] Helmholtz H. Zur Theorie der stationären Ströme in reibenden Flüssigkeiten. *Verh. Naturh.-Med. Ver. Heidelberg*, **11**:223, 1868.

[125] Helmholtz H. Über die physikalische Bedeutung des Prinicips der kleinsten Wirkung (Fortsetzung). *J. f. d. reine u. angewandte Math.*, **100**:213–222, 1887.

[126] Henson V. E., and Vassilevski P. S. Element-free AMGe: General algorithms for computing interpolation weights in AMG. *SIAM J. Sci. Comput.*, **23**(2):629–650, 2001.

[127] Herrmann D. Implementierung eines Ritzschen Verfahrens zur Berechnung von fluiden Filmen über Topografien. Bachelor's thesis, Heilbronn University, 2018.

[128] Hirsch A. Über eine charakteristische Eigenschaft der Differentialgleichungen der Variationsrechnung. *Mathematische Annalen*, **49**(1):49–72, 1897.

[129] Horner M., Metcalfe G., Wiggins S., and Ottino J. Transport enhancement mechanisms in open cavities. *J. Fluid Mech.*, **452**:199–229, 2002.

[130] Howarth L. The boundary layer in three dimensional flow – Part II. The flow near a stagnation point. *The London, Edinburgh, and Dublin Phil. Mag. J. Sci.*, **42** (335):1433–1440, 1951.

[131] Howison S. D. Complex variable methods in Hele-Shaw moving boundary problems. *Europ. J. Appl. Math.*, **3**:209–224, 1992.

[132] Hutter K., Svendsen B., and Rickenmann D. Debris flow modeling: A review. *Continuum Mechanics and Thermodynamics*, **8**(1):1–35, 1994.

[133] Jackson J. D. *Classical Electrodynamics*. John Wiley & Sons Inc., New York, 1999.

[134] Jackson J. D., and Okun L. B. Historical roots of gauge invariance. *Rev. Mod. Phys.*, **73**:663–680, Sep 2001. doi: 10.1103/RevModPhys.73.663.

[135] Jiang B. N., Lin T. L., and Povinelli L. A. Large-scale computation of incompressible viscous flow by least-squares finite element method. *Comput. Methods Appl. Mech. Engrg.*, **114**:213–231, 1994.

[136] John L., Pustejovska P., Wohlmuth B., and Rüde U. Energy-corrected finite element methods for the Stokes system. *IMA J. Num. Analysis*, **37**(2):687–729, 2016.

[137] John V. *Finite Element Methods for Incompressible Flow Problems*. Springer Series in Computational Mathematics. Springer International Publishing, 2016.

[138] Johnston J. Z., and Tabarrok B. Stream function - stress function approach to incompressible flows. In Taylor C., Morgan K., and Brebbia C. A., editors, *Numerical Methods in Laminar and Turbulent Flow*, pages 81–88, 1978.

[139] Jones R. L., and Stern K. H. *Metallurgical and Ceramic Protective Coatings*. Chapman Hall, London, 1996.

[140] Joseph D. D. Potential flow of viscous fluids: Historical notes. *Int. J. Multiphase Flow*, **32**(3):285–310, 2006.

[141] Kaku M. *Quantum Field Theory: A Modern Introduction*. Oxford Univ. Press, 1993.

[142] Kambe T. Variational formulation of ideal fluid flows according to gauge principle. *Fluid Dynamics Research*, **40**(6):399–426, 2008.

[143] Kambe T. On fluid Maxwell equations. In *Frontiers of Fundamental Physics and Physics Education Research*, pages 287–295. Springer, 2014.

[144] Karageorghis A., and Fairweather G. The method of fundamental solutions for the numerical solution of the biharmonic equation. *J. Comput. Phys.*, **69**(2): 434–459, 1987.

[145] Kaufman A. N. Dissipative Hamiltonian systems: a unifying principle. *Physics Letters A*, **100**(8):419–422, 1983.

[146] Kerswell R. Variational principle for the Navier-Stokes equations. *Physical Review E*, **59**(5):5482, 1999.

[147] Kistler S. F., and Schweizer P. M. *Liquid Film Coating*. Chapman Hall, New York, 1997.

[148] Koide T., and Kodama T. Navier–Stokes equation by stochastic variational method. *arXiv preprint arXiv:1105.6256*, 2001.

[149] Kondrat'ev V. A. Boundary value problems for elliptic equations in domains with conical or angular points. *Trudy Moskovskogo Matematicheskogo Obshchestva*, **16**: 209–292, 1967.

[150] Kotulek J. Historical notes on the inverse problem of the calculus of variations. Master's thesis, Silesian University in Opava, 2002.

[151] Kozlov V. A., Mazia V., and Rossmann J. *Elliptic boundary value problems in domains with point singularities*, volume 52. American Math. Soc., 1997.

[152] Krechel A., and Stüben K. Parallel algebraic multigrid based on subdomain blocking. *Parallel Computing*, **27**:1009–1031, 2001.

[153] Ku H. C., Hirsh R. S., and Taylor T. D. A pseudospectral method for solution of the three-dimensional incompressible Navier-Stokes equations. *J. Comput. Phys.*, **70**:439–462, 1987.

[154] Lamb H. *Hydrodynamics*. Cambridge University Press, 6th edition, 1974.

[155] Landau L. D. *The Classical Theory of Fields*, volume 2. Elsevier, 2013.

[156] Landau L. D., and Lifschitz E. M. *Fluid Mechanics*. Pergamon, 2nd edition, 1987.

[157] Lee E., Manteuffel T. A., and Westphal C. Weighted-norm first-order system least squares (FOSLS) for problems with corner singularities. *SIAM J. Num. Analysis*, **44**(5):1974–1996, 2006.

[158] Lee Y., Thompson H., and Gaskell P. Three-dimensional thin film and droplet flows over and past surface features with complex physics. *Computers & Fluids*, **46** (1):306–311, 2011.

[159] Legendre R. Solutions plus complète du problème Blasius. *Comptes Rendus*, **228**: 2008–2010, 1949.

[160] Lenoir M. Optimal isoparametric finite elements and error estimates for domains involving curved boundaries. *SIAM J. Num. Analysis*, **23**(3):562–580, 1986.

[161] Lew A., Marsden J. E., Ortiz M., and West M. An overview of variational integrators. In *Finite Element Methods: 1970's and Beyond*. International Center for Numerical Methods in Engineering (CIMNE), 2004.

[162] Lin C. C. Hydrodynamics of Helium II. In *Proc. Int. School of Physics "Enrico Fermi"*, volume 21, New York, 1963. Academic Press.

[163] Lions J., and Magenes E. *Non-Homogenous Boundary Value Problems and Applications I*, volume 181 of *Die Grundlehren der mathematishen Wissenschaften in Einzeldarstellungen*. Springer-Verlag, Berlin, 1972.

[164] Lubuma J. M. S., and Nicaise S. Dirichlet problems in polyhedral domains II: Approximation by FEM and BEM. *J. Comput. Appl. Math.*, **61**(1):13–27, 1995.

[165] Luca I., Hutter K., Tai Y. C., and Kuo C. Y. A hierarchy of avalanche models on arbitrary topography. *Acta Mechanica*, **205**(1-4):121–149, 2009.

[166] Luchini P. Higher-order difference approximations of the Navier-Stokes equations. *Int. J. Num. Meth. Fluids*, **12**(5):491–506, 1991.

[167] Luke J. A variational principle for a fluid with a free surface. *J. Fluid Mech.*, **27** (2):395–397, 1967.

[168] Madelung E. Quantentheorie in hydrodynamischer Form. *Zeitschrift für Physik*, **40**(3):322–326, 1927. ISSN 0044-3328.

[169] Magri F. Variational formulation for every linear equation. *Int. J. Engineering Sci.*, **12**(6):537–549, 1974.

[170] Malevich A. E., Mityushev V. V., and Adler P. M. Couette flow in channels with wavy walls. *Acta Mechanica*, **197**:247–283, 2008.

[171] Marner F., Gaskell P. H., and Scholle M. On a potential-velocity formulation of Navier-Stokes equations. *Physical Mesomechanics*, **17**(4):341–348, 2014.

[172] Marner F., Gaskell P. H., and Scholle M. A complex-valued first integral of Navier-Stokes equations: Unsteady Couette flow in a corrugated channel system. *J. Math. Phys.*, **58**(4):043102, 2017.

[173] Marner F., Scholle M., Herrmann D., and Gaskell P. H. Competing Lagrangians for incompressible and compressible viscous flow. *Roy. Soc. Open Sci.*, **6** (1):181595, 2019.

[174] MathWorks. Partial differential equation toolbox, user's guide. Technical report, The MathWorks Inc., 3 Apple Hill Drive, Natick, MA 01760-2098, 2017.

[175] Mayes C., Schlichting H., Krause E., Oertel H. J., and Gersten K. *Boundary-Layer Theory*. Physic and Astronomy. Springer, 2003. ISBN 9783540662709.

[176] Mellmann M. Implementierung eines neuartigen iterativen Lösers für die Simulation der Filmströmung über geneigte, profilierte Substrate. Bachelor's thesis, Heilbronn University, 2017.

[177] Mikhlin S. G. *Integral Equations and their Applications to certain Problems in Mechanics, Mathematical Physics and Technology.* Pergamon Press, New York, 1957.

[178] Millikan C. B. On the steady motion of viscous, incompressible fluids; with particular reference to a variation principle. *The London, Edinburgh, and Dublin Phil. Mag. J. Sci.*, **7**(44):641–662, 1929.

[179] Mobbs S. Variational principles for perfect and dissipative fluid flows. *Proc. R. Soc. Lond. A*, **381**(1781):457–468, 1982.

[180] Moffatt H. K. Viscous and resistive eddies near a sharp corner. *J. Fluid Mech.*, **18**(1):1–18, 1964.

[181] Montero R. S., Llorente I. M., and Salas M. D. Robust multigrid algorithms for the Navier–Stokes equations. *J. Comp. Phys.*, **173**(2):412–432, 2001.

[182] Morandi G., Ferrario C., Vecchio G. L., Marmo G., and Rubano C. The inverse problem in the calculus of variations and the geometry of the tangent bundle. *Physics Reports*, **188**(3-4):147–284, 1990.

[183] Morrison P., Francoise J., Naber G., and Tsou S. Hamiltonian fluid dynamics. *Encyclopedia of Math. Phys.*, pages 593–600, 2006.

[184] Muskhelishvili N. I. *Some Basic Problems of the Mathematical Theory of Elasticity.* Noordhoff, Groningen, 1953.

[185] Neuber H. Ein neuer Ansatz zur Lösung räumlicher Probleme der Elastizitätstheorie. *J. Appl. Math. Mech.*, **14**:2008–2010, 1934.

[186] Nguyen P. K., and Bontozoglou V. Steady solutions of inertial film flow along strongly undulated substrates. *Phys. Fluids*, **23**(5):052103, 2011.

[187] Ockendon H., and Ockendon J. R. *Viscous Flow, Cambridge Texts in Applied Mathematics.* Cambridge University Press, Cambridge, 1995.

[188] Olsson P. *Transport Phenomena in Newtonian Fluids - A Concise Primer.* SpringerBriefs in Applied Sciences and Technology. Springer International Publishing, 2013. ISBN 9783319013091.

[189] Onsager L. The motion of ions: principles and concepts. In *The Collected Works Of Lars Onsager: (With Commentary)*, pages 46–60. World Scientific, 1996.

[190] Owens R. G., and Phillips T. N. Mass- and momentum-conserving spectral methods for Stokes flow. *J. Comput. Appl. Math.*, **53**:185–206, 1994.

[191] Panton R. L. *Incompressible Flow.* John Wiley & Sons Inc., 1996.

[192] Payette G. S., and Reddy J. N. On the roles of minimization and linearization in least-squares finite element models of nonlinear boundary value problems. *Electr. J. Differential Equations*, **202**:1–7, 2012.

[193] Pollak T., and Aksel N. Crucial flow stabilization and multiple instability branches of gravity-driven films over topography. *Phys. Fluids*, **25**, 2013.

[194] Pozrikidis C. *Boundary Integral and Singularity Methods for Linearized Viscous Flow*. Cambridge University Press, 1992.

[195] Prakash J., Lavrenteva O. M., and Nir A. Application of Clebsch variables to fluid-body interaction in presence of non-uniform vorticity. *Phys. Fluids*, **26**(7): 077102, 2014.

[196] Prigogine I., and Glansdorff P. Variational properties and fluctuation theory. *Physica*, **31**(8):1242–1256, 1965.

[197] Ramm A. G., and Fabrikant V. I. Justification of Fabrikant's method for solving mixed problems of potential theory. *Computers Math. Applic.*, **22**(6):97–103, 1991.

[198] Ranger K. Parametrization of the two and three-dimensional motion of a viscous incompressible liquid. *Quarterly J. Appl. Math.*, **64**(3):401–412, 2006.

[199] Ranger K. B. Parametrization of general solutions for the Navier-Stokes equations. *Quarterly J. Appl. Math.*, **52**:335–341, 1994.

[200] Rayleigh O. M. On the motion of a viscous fluid. *The London, Edinburgh, and Dublin Phil. Mag. J. Sci.*, **26**(154):776–786, 1913.

[201] Reddiger M. The Madelung picture as a foundation of geometric quantum theory. *Foundations of Physics*, **47**(10):1317–1367, 2017.

[202] Richardson S. Two-dimensional slow viscous flows with time-dependent free boundaries driven by surface tension. *Europ. J. Appl. Math.*, **3**:193–207, 1992.

[203] Richter A. Test einer Variationsformulierung für die Strömungssimulation. Master's thesis, Heilbronn University, 2014.

[204] Riewe F. Nonconservative Lagrangian and Hamiltonian mechanics. *Physical Review E*, **53**(2):1890, 1996.

[205] Rogula D. Continuum field theory of string-like objects. Dislocations and superconducting vortices. *ZAMP*, **57**(1):123–132, 2005. ISSN 0044-2275.

[206] Roitberg B. Y., and Roitberg Y. A. Elliptic boundary-value problems in nonsmooth domains. *Ukrainian Mathematical Journal*, **47**(5):808–817, 1995.

[207] Roitberg Y. A., and Seftel Z. A theorem on homeomorphisms for elliptic systems and its applications. *Math. USSR Sbornik*, **7**:439–465, 1969.

[208] Rosen P. On variational principles for irreversible processes. *J. Chemical Phys.*, **21**(7):1220–1221, 1953.

[209] Rudge J. F. Analytical solutions of compacting flow past a sphere. *J. Fluid Mech.*, **746**:466–497, 2014.

[210] Ruge J. W., and Stüben K. Algebraic multigrid. In McCormick S. F., editor, *Multigrid Methods*. SIAM, Philadelphia, 2003.

[211] Saad Y. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, 2nd edition, 2003.

[212] Sala M., Tuminaro R. S., Hu J. J., and Gee M. W. Ml 4.0 smoothed aggregation user's guide. Technical report, Sandia National Laboratories, 2005.

[213] Salmon R. Hamiltonian fluid mechanics. *Ann. Rev. Fluid Mech.*, **20**(1):225–256, 1988.

[214] Santilli R. M. *Foundations of Theoretical Mechanics I. The Inverse Problem in Newtonian Mechanics*. Springer, 1978.

[215] Saranent J., and Wendland W. A least squares approximation method for first-order elliptic systems of plane. *Applicable Analysis*, **14**(1):27–42, 1982.

[216] Scholle M. *Das Hamiltonsche Prinzip in der Kontinuumstheorie nichtdissipativer und dissipativer Systeme. Ein neues Konzept zur Konstruktion von Lagrangedichten.* PhD thesis, University of Paderborn, 1999.

[217] Scholle M. Creeping Couette flow over an undulated plate. *Arch. Appl. Mech.*, **73**:823–840, 2004.

[218] Scholle M. Construction of Lagrangians in continuum theories. *Proc. R. Soc. Lond. A*, **460**:3241–3260, 2004.

[219] Scholle M. Einfluss der Randgeometrie auf die Strömung in fluiden Schichten. Habilitationsschrift, Universität Bayreuth, 2004.

[220] Scholle M. Comment on "Variational approach to the volume viscosity of fluids" [Phys. Fluids 18, 047101 (2006)]. *Phys. Fluids*, **18**(10):109101, 2006.

[221] Scholle M. Hydrodynamical modelling of lubricant friction between rough surfaces. *Tribology International*, **40**(6):1004 – 1011, 2007. ISSN 0301-679X.

[222] Scholle M. Variational formulations for viscous flow. *PAMM*, **14**(1):611–612, 2014. ISSN 1617-7061.

[223] Scholle M. A variational approach for viscous flow. In Ott P., editor, *Proceedings of the Heilbronn Workshop on Research in Mechatronics*, pages 1–8. Heilbronn University, 2014.

[224] Scholle M., and Anthony K. H. Line-shaped objects and their balances related with gauge symmetries in continuum theories. *Proc. R. Soc. Lond. A*, **460**:875–896, 2004.

[225] Scholle M., and Marner F. A generalized Clebsch transformation leading to a first integral of Navier–Stokes equations. *Physics Letters A*, **380**(40):3258–3261, 2016.

[226] Scholle M., and Marner F. A non-conventional discontinuous Lagrangian for viscous flow. *Roy. Soc. Open Sci.*, **4**(2), 2017.

[227] Scholle M., Wierschem A., and Aksel N. Creeping films with vortices over strongly undulated bottoms. *Acta Mechanica*, **168**(3):167–193, 2004. ISSN 1619-6937.

[228] Scholle M., Haas A., Aksel N., Wilson M. C. T., Thompson H. M., and Gaskell P. Competing geometric and inertial effects on local flow structure in thick gravity-driven fluid films. *Phys. Fluids*, **20**(12), 2008.

[229] Scholle M., Haas A., Aksel N., Wilson M. C. T., Thompson H. M., and Gaskell P. H. Eddy genesis and manipulation in plane laminar shear flow. *Phys. Fluids*, **21**:073602, 2009.

[230] Scholle M., Haas A., and Gaskell P. H. A first integral of Navier-Stokes equations and its applications. *Proc. R. Soc. Lond. A*, **467**:127–143, 2011.

[231] Scholle M., Gaskell P. H., and Marner F. Exact integration of the unsteady incompressible Navier-Stokes equations, gauge criteria, and applications. *J. Math. Phys.*, **59**(4):043101, 2018. doi: 10.1063/1.5031119.

[232] Scholle M., Gaskell P. H., and Marner F. A potential field description for gravity-driven film flow over piece-wise planar topography. *Fluids*, **4**(2):4020082, 2019.

[233] Schönberg M. Vortex motions of the Madelung fluid. *Il Nuovo Cimento*, **1**(4): 543–580, 1955.

[234] Schörner M., Reck D., and Aksel N. Does the topography's specific shape matter in general for the stability of film flows? *Phys. Fluids*, **27**(4):042103, 2015.

[235] Schwarzbach Y. *The Noether Theorems – Invariance and Conservation Laws in the 20th Century*. Springer, 2010.

[236] Seliger R., and Witham G. B. Variational principles in continuum mechanics. *Proc. R. Soc., London*, **A305**:1–25, 1968. doi: 10.1098/rspa.1968.0103.

[237] Shewchuk J. R. An introduction to the conjugate gradient method without the agonizing pain. School of Computer Science, Carnegie Mellon University, Pittsburgh, August 1994.

[238] Slade D., Veremieiev S., Lee Y., and Gaskell P. Gravity-driven thin film flow: The influence of topography and surface tension gradient on rivulet formation. *Chemical Engineering and Processing: Process Intensification*, **68**:7–12, 2013.

[239] Slade D. R. J. *Gravity-Driven Thin Liquid Films: Rivulets and Flow Dynamics*. PhD thesis, University of Leeds, 2013.

[240] Sommerfeld A. *Mechanik der deformierbaren Medien*. Vorlesungen über Theoretische Physik. Harri Deutsch, 6th edition, 1992.

[241] Spurk H., and Aksel N. *Fluid Mechanics*. Springer Berlin Heidelberg, 2nd edition, 2008.

[242] Stüben K. An introduction to algebraic multigrid. In Trottenberg U., Oosterlee C. W. and Schüller A., editor, *Multigrid*. Academic Press, San Diego, CA, 2001.

[243] Strang G., and Fix G. J. *An Analysis of the Finite Element Method*. Wellesley-Cambridge Press, 2008.

[244] Stück A. *Adjoint Navier-Stokes Methods for Hydrodynamic Shape Optimisation*. PhD thesis, Technical University of Hamburg, 2012.

[245] Szulczewska B., Zbicinski I., and Gorak A. Liquid flow on structured packing: CFD simulation and experimental study. *Chemical Engineering & Technology*, **26** (5):580–584, 2003.

[246] Takagi D., and Huppert H. E. Flow and instability of thin films on a sphere. *J. Fluid Mech.*, **647**:221–238, 2009.

[247] Tanveer S., and Vasconcelos G. L. Time-evolving bubbles in two dimensional Stokes flow. *J. Fluid Mech.*, **301**:325–344, 1995.

[248] Tanzosh J. P., and Stone H. A. A general approach for analyzing the arbitrary motion of a circular disk in a Stokes flow. *Chem. Eng. Comm.*, **148-150**:333–346, 1996.

[249] Taylor C. M. *Engine Tribology*, volume 26. Elsevier, 1993.

[250] Tezduyar T., Glowinski R., and Liou J. Petrov-Galerkin methods on multiply connected domains for the vorticity-stream function formulation of the incompressible Navier-Stokes equations. *Int. J. Num. Meth. Fluids*, **8**(10):1269–1290, 1988.

[251] Thatcher R. A least squares method for solving biharmonic problems. *SIAM J. Num. Analysis*, **38**(5):1523–1539, 2000.

[252] Thatcher R. W. A least squares method for Stokes flow based on stress and stream functions. Technical report, Manchester Centre for Computational Mathematics, Manchester University / UMIST, 1998.

[253] Tonti E. On the variational formulation for linear initial value problems. *Annali di Matematica pura ed applicata*, **95**(1):331–359, 1973.

[254] Tonti E. A general solution of the inverse problem of the calculus of variations. *Hadronic Journal*, **5**:1404–1450, 1982.

[255] Tonti E. Variational formulation for every nonlinear problem. *Int. J. Engineering Sci.*, **22**(11-12):1343–1371, 1984.

[256] Tran-Cong T., and Blake J. R. General solutions of the Stokes' flow equations. *J. Math. Anal. Appl.*, **90**(1):72–84, 1982.

[257] Trifonov Y. Stability of a film flowing down an inclined corrugated plate: The direct Navier-Stokes computations and Floquet theory. *Phys. Fluids*, **26**(11):114101, 2014.

[258] Trifonov Y. Y. Viscous liquid film flows over a periodic surface. *Int. J. Multiphase Flow*, **24**(7):1139–1161, 1999.

[259] Trottenberg U., Oosterlee C., and Schüller A. *Multigrid*. Academic Press, San Diego, CA, 2001.

[260] Tuckerman L. S. Divergence-free velocity fields in nonperiodic geometries. *J. Comput. Phys.*, **80**(2):403–441, 1989. ISSN 0021-9991.

[261] Vainberg M. *Variational Methods for the Study of Nonlinear Equations*. Holden-Day, San Francisco, 1964.

[262] van Dantzig D. On the phenomenological thermodynamics of moving matter. *Physica*, **6**(7):673–704, 1939. ISSN 0031-8914.

[263] Vanek P., Mandel J., and Brezina M. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing*, **56**:179–196, 1996.

[264] Venturi D. Conjugate flow action functionals. *J. Math. Phys.*, **54**(11):113502, 2013.

[265] Veremieiev S. *Gravity-Driven Continuous Thin Film Flow over Topography*. PhD thesis, University of Leeds, 2011.

[266] Veremieiev S., Thompson H. M., Lee Y. C., and Gaskell P. H. Computing inertial two- and three-dimensional thin film flow on planar surfaces featuring topography. *Computers & Fluids*, **39(3)**:431–450, 2010.

[267] Veremieiev S., Thompson H. M., and Gaskell P. H. Free-surface film flow over topography: Full three-dimensional finite element solutions. *Computers & Fluids*, **122**:66–82, 2015.

[268] Vlasogiannis P., Karagiannis G., Argyropoulos P., and Bontozoglou V. Air–water two-phase flow and heat transfer in a plate heat exchanger. *Int. J. Multiphase Flow*, **28**(5):757–772, 2002.

[269] Volevich L. R. Solubility of boundary value problems for general elliptic systems. *Matematicheskii Sbornik*, **110**(3):373–416, 1965.

[270] Volterra V. *Leçons sur les fonctions de lignes*. Goutier-Villard, Paris, 1913.

[271] Wagner H. J. On the use of Clebsch potentials in the Lagrangian formulation of classical electrodynamics. *Physics Letters*, **A292**:246–250, 2002.

[272] Wang Y., Shu C., Teo C. J., Wu J., and Yang L. Three-dimensional lattice Boltzmann flux solver and its applications to incompressible isothermal and thermal flows. *Com. Comput. Phys.*, **18**:593–620, 2015.

[273] Webb R. *Principles of Enhanced Heat Transfer, 1994.* John Wiley & Sons, New York, 1994.

[274] Weinstein S. J., and Ruschak K. J. Three-dimensional lattice Boltzmann flux solver and its applications to incompressible isothermal and thermal flows. *Annu. Rev. Fluid Mech.*, **36**:29–53, 2004.

[275] Wendland W. L. *Elliptic Systems in the Plane.* Monographs and Studies in Mathematics. Pitman, London, San Francisco, 1979.

[276] Wierschem A., and Aksel N. Influence of inertia on eddies created in films creeping over strongly undulated substrates. *Phys. Fluids*, **16**(12):4566–4574, 2004.

[277] Wierschem A., Lepski C., and Aksel N. Effect of long undulated bottoms on thin gravity-driven films. *Acta Mechanica*, **179**(1-2):41–66, 2005.

[278] Wierschem A., Pollak T., Heining C., and Aksel N. Suppression of eddies in films over topography. *Phys. Fluids*, **22**(11):113603, 2010.

[279] Wilson M. C. T., Summers J. L., Kapur N., and Gaskell P. H. Stirring and transport enhancement in a continuously modulated free-surface flow. *J. Fluid Mech.*, **565**:319–351, 2006.

[280] Wong V. W., and Tung S. C. Overview of automotive engine friction and reduction trends–effects of surface, material, and lubricant-additive technologies. *Friction*, **4**(1):1–28, 2016.

[281] Woodhouse F. G., and Goldstein R. E. Shear-driven circulation patterns in lipid membrane vesicles. *J. Fluid Mech.*, **705**:165–175, 2012.

[282] Wright N. G., and Gaskell P. H. An efficient multigrid approach to solving highly recirculating flows. *Computers & Fluids*, **24**(1):63–79, 1995.

[283] Wu A., and Yang C. N. Evolution of the concept of the vector potential in the description of fundamental interactions. *Int. J. Modern Physics A*, **21**(16):3235–3277, 2006.

[284] Yang C. N. The conceptual origins of and gauge theory. *Physics Today*, **67**(11):45, 2014.

[285] Yasue K. A variational principle for the Navier-Stokes equation. *J. Funct. Analysis*, **51**(2):133–141, 1983.

[286] Yoshida Z. Clebsch parameterization: Basic properties and remarks on its applications. *J. Math. Phys.*, **50**(11):113101, 2009.

[287] Zuckerwar A. J., and Ash R. L. Response to "Comment on 'Variational approach to the volume viscosity of fluids"' [Phys. Fluids 18, 109101 (2006)]. *Phys. Fluids*, **18**(10):109102, 2006.

[288] Zuckerwar A. J., and Ash R. L. Volume viscosity in fluids with multiple dissipative processes. *Phys. Fluids*, **21**:1091–1106, 2009.