Lecture Notes in Geosystems Mathematics and Computing

Matthias Albert Augustin

A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs

Birkhäuser

Lecture Notes in Geosystems Mathematics and Computing

Birkhäuser

More information about this series at
http://birkhauser-science.com/series/13390

Matthias Albert Augustin

A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs

Birkhäuser

Matthias Albert Augustin
AG Geomathematik
Technische Universität Kaiserslautern
Kaiserslautern
Germany

Geosystems Mathematics
ISBN 978-3-319-17078-7 ISBN 978-3-319-17079-4 (eBook)
DOI 10.1007/978-3-319-17079-4

Library of Congress Control Number: 2015942919

Mathematics Subject Classification (2010): 35Qxx, 65-XX

Springer Cham Heidelberg New York Dordrecht London

Cover design: deblik, Berlin

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.birkhauser-science.com)

There is no branch of mathematics, however abstract, which may not some day be applied to phenomena of the real world.

Nicolai Ivanovich Lobachevsky, [237]

This work was submitted as PhD thesis at the University of Kaiserslautern, Department of Mathematics and fully accepted.

Acknowledgments

First of all, I want to express my gratitude to Prof. Dr. Willi Freeden for giving me the possibility to work on an interesting topic in a field full of mathematical challenges and with a high applicability to real-world problems. I have always appreciated his support during every step of this PhD project, his many helpful and valuable suggestions, and his confidence in my abilities.

Second, I would like to thank Prof. Dr. Thomas Sonar for acting as my second supervisor and his interest in my work.

Special thanks go to Dipl.–Math./M.Sc. C. Blick, Dipl.–Math. S. Eberle, Dr. C. Gerhards, Dr. M. Gutting, Dipl.–Math. S. Möhringer, Dr. H. Nutz, Dr. I. Ostermann, and Dipl.–Phys. C. Thome for their great support, several fruitful discussions, helpful suggestions, and proofreading my thesis.

Moreover, I am grateful to all former and present members of the Geomathematics Group Kaiserslautern for the productive and pleasant atmosphere during the whole time.

Finally, I would like to thank the "Fraunhofer Institute for Industrial Mathematics ITWM" ("Institut für Techno- und Wirtschaftmathematik") for financial support, the "German National Academic Foundation" ("Studienstiftung des deutschen Volkes") for giving me a scholarship during my work on this PhD–thesis and financial support during my prior studies, as well as the "Center for Mathematics and Computational Modeling $(CM)^2$" and the "German Academic Exchange Service" ("Deutscher Akademischer Austauschdienst DAAD") for financial support which allowed me to present my work at the Joint Mathematics Meeting of the American Mathematical Society (AMS) and the Mathematical Association of America (MAA) in Boston (2012) and San Diego (2013), respectively.

Last but not least, I want to say "thank you" to Julia for her support, her patience, and for filling my life with love and joy.

Contents

Chapter 1
Introduction

1.1 Motivation: Geothermal Background

The human hunger for energy is steadily increasing. Satisfying this hunger by fossil fuels like coal, oil, or gas becomes more and more difficult. On the one hand, this is because of shrinking resources. On the other hand, the usage of fossil fuels has to be reduced in order to prevent further climate changes to which the CO_2 production of facilities using fossil fuels contributes heavily. Often, nuclear power plants are mentioned as an alternative. Unfortunately, instead of the danger of climate change, other problems arise when using nuclear power. One is the risk of a maximum credible accident (MCA). Although engineers and scientists planning and operating nuclear power plants try to minimize this risk, it will never vanish. As the incidents at Three Mile Island, USA, in 1979, at Chernobyl, Ukrainian SSR (now Ukraine), in 1986, or at Fukushima, Japan, in 2011 clearly demonstrated, accidents involving nuclear power plants can be fatal for people and environment. The bigger problem is the radioactive waste produced during energy production. Until today, it seems that nobody knows how to store this waste safely and there is no known final depot in any country using nuclear power. Nevertheless, it seems that nuclear power plants are indispensable as a transition technology from energy production using fossil fuels to power plants operating on renewable resources, such as hydropower, wind, or solar energy.

In the context of energy based on wind or solar power, it is often objected that those show characteristic fluctuations on a daily and annual scale as they depend on changes in weather and season [117, 273]. For this reason, they are unable to provide a steady baseload. Moreover, power plants operating on these resources use large amounts of land, therefore having a visual impact on scenery, with wind farms also shadowing parts of the landscape and being reported to generate significant amounts of noise [117, 273].

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4_1

Table 1.1 Technical potential of energy sources (according to [117])

Resource	EJ/year	$\approx 10^3$TWh/year
Hydropower	50	13.9
Biomass	276	76.7
Wind energy	640	177.8
Solar energy	1,575	437.5
Geothermal energy	5,000	1,388.9
Total	7,600	2,111.1

One rather important, but unfortunately in public opinion underestimated renewable resource is the heat stored within the Earth's crust which can be used by geothermal facilities. The direct use of geothermal energy can be dated back to historical times all over the world, e.g., Romans, Greeks, Indians, or Japanese, whereas electricity production from geothermal energy began in Larderello, Italy, in 1913 after first experiments were conducted there in 1904. For a further historical outline, see, e.g., [23, 99].

According to the World Energy Assessment [117], geothermal energy has the highest potential capacity of all renewables (see Table 1.1). Moreover, as opposed to most other renewables, geothermal energy does not depend on external influences as certain weather conditions and is, thus, not influenced by daily or seasonal changes in weather or climate. This makes it highly reliable and guarantees a stable supply which allows to install geothermal energy as a baseload provider [100].

Despite this high potential, only 13.7 % of the world's primary energy demands of 418 EJ (≈116,111.1 TWh) in 2001 was provided by renewables [118]. The installed capacity of geothermal energy was 48,493 MW for direct usage, providing 0.42383 EJ (≈117.7 TWh) in 78 countries in 2009 [183], and 10,898 MW for electric power, providing 67.246 TWh (≈0.24209 EJ) in electricity in 24 countries in 2010 [33]. Thus, geothermal energy production is a field with an extremely high growing potential. For an overview of the geothermal potential in Germany, see [146, 220, 249, 250] and the references therein.

A critical quantity for the use of geothermal energy in industrial applications is the temperature. For direct use, i.e., heating applications, temperatures of 100–150 °C are sufficient, whereas 150–200 °C are needed for electricity production [273]. Assuming that temperature increases with depth typically by 30 °C per kilometer [103] and an average surface temperature of 15 °C [117], this means that using geothermal energy requires to access geothermal resources at depths between 3 and 7 km. Fortunately, the geothermal gradient varies widely and can be as large as 50 °C [273] or even 150 °C per kilometer [103].

There are many different kinds of classification for geothermal reservoirs. The following classification is taken verbatim from [241]:

1. *Volcanic geothermal systems* are in one way or another associated with volcanic activity. The heat sources for such systems are hot intrusions or magma. They are most often situated inside, or close to, volcanic complexes such as calderas, most of them at plate boundaries, but some in hot spot areas. Permeable fractures and fault zones mostly control the flow of water in volcanic systems.

2. In *convective fracture controlled systems* the heat source is the hot crust at depth in tectonically active areas with above average heat flow. Here the geothermal water has circulated to considerable depth (>1 km), through mostly vertical fractures, to mine the heat from the rocks.
3. *Sedimentary geothermal systems* are found in many of the major sedimentary basins of the world. These systems owe their existence to the occurrence of permeable sedimentary layers at great depths (>1 km) and above average geothermal gradients (>30 °C/km). These systems are conductive in nature rather than convective, even though fractures and faults play a role in some cases. Some convective systems (2) may, however, be embedded in sedimentary rocks.
4. *Geo-pressured systems* are analogous to geo-pressured oil or gas reservoirs, where fluid caught in stratigraphic traps may have pressure close to lithostatic values. Such systems are generally fairly deep; hence, they are categorized as geothermal.
5. *Hot dry rock (HDR) or enhanced (engineered) geothermal systems (EGS)* consist of volumes of rock that have been heated to useful temperatures by volcanism or abnormally high heat flow, but have low permeability or are virtually impermeable. Therefore, they cannot be exploited in a conventional way. However, experiments have been conducted in a number of locations to use hydro-fracturing to try to create artificial reservoirs in such systems, or to enhance already existent fracture systems. Such systems will mostly be used through production/re-injection doublets.
6. *Shallow resources* refer to the normal heat flux through near surface formations and the thermal energy stored in the rocks and warm groundwater systems near the surface of the Earth's crust. Recent developments in the application of ground source heat pumps have opened up a new dimension in utilizing these resources.

The water-bearing layers of permeable, porous rock or sediments are also called aquifers. Classes 1–4 are also called hydrothermal systems, as they naturally contain water, whereas HDR systems are also called petrothermal systems (see also Fig. 1.1). Hot wet rock (HWR) systems can be seen as intermediate between HDR and hydrothermal systems, as the reservoir contains water, but can only be used after the artificial creation of fractures [23]. It is, thus, also a petrothermal system. Another type of geothermal systems are deep heat exchangers or thermowells [161]. They consist of a coaxial pipe, therefore using only one borehole, through which water circulates to extract the Earth's heat from about 2 km depth. Unfortunately, the advantage of having a closed system is counteracted by rather low productivity.

The systems mentioned above, with the possible exception of shallow resources, are so-called deep geothermal systems. There is also so-called near-surface geothermal energy, which covers all systems with a depth of 400 m or less. Usually, this term is used for all kinds of geothermal collectors or thermo-active pipes, primarily used for heating or energy production in private homes. In this thesis, we are only concerned with deep geothermal systems.

A classification of hydrothermal reservoirs by temperature may yield up to seven classes [243], whereas the physical state of the water in the reservoir – primarily liquid, primarily steam/vapor, or a two-phase mixture – should also be taken into account (see [241] and the references therein). Moreover, geothermal systems can be classified by enthalpy, i.e., the energy density of the reservoir fluids given in kJ/kg, distinguishing low-enthalpy systems with less than 800 kJ/kg and high-enthalpy systems with more than 800 kJ/kg [241]. As permeability and, thus, fluid

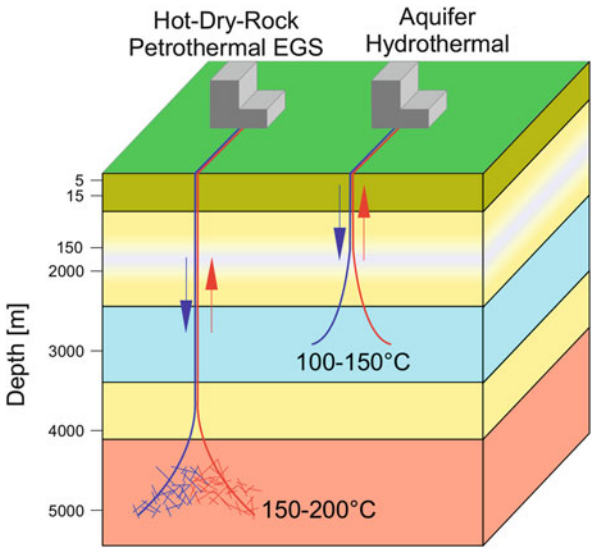

Fig. 1.1 Examples for different types of geothermal facilities (not to scale). On the left-hand side, an EGS is shown with deep drilled wells into hot dry rock (*light red*) using an artificial fracture system. On the right-hand side, a sedimentary geothermal system can be seen, using hot water in a natural aquifer (turquoise)

temperature in petrothermal systems depend on the artificially created fracture pattern, classification of such systems by temperature does not yield valuable information.

For further information on geothermal energy, the reader is referred to the already mentioned references and, e.g., [23, 66, 126] and the references therein.

As with all technological progress, geothermal energy production and usage has to meet certain challenges and risks. The following list is compiled from [23, 103].

Geothermal resources have a fixed location and cannot be moved, thus, restricting the choice of locations for geothermal power plants. Moreover, as promising geothermal fields are often found in remote areas, new infrastructure may have to be build. Thus, issues of land ownership or right-of-way for roads or transmission lines arise.

Geothermal fluids may contain non-condensable gases like hydrogen sulfide, carbon dioxide, or methane which, if released, are considered harmful to the environment but have to be removed before steam enters a turbine. An option to meet this challenge is the usage of binary cycles, in which the withdrawn geothermal fluid is used to heat another, non-contaminated fluid in a separate flow cycle and re-injected without releasing non-condensable gases.

Exploring geothermal reservoirs couples a high upfront risk with a need of high upfront investment and a long project development cycle, resulting in the discouragement of potential investors.

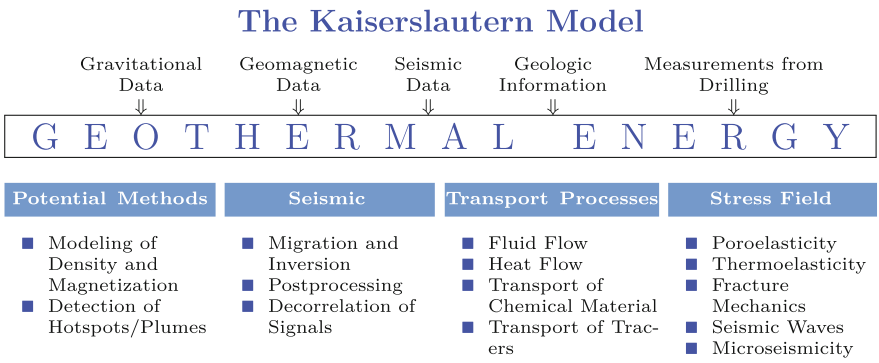

Fig. 1.2 Column model characterizing deep geothermal systems (as developed by the Geomathematics Group at the University of Kaiserslautern)

Extracting geothermal fluids from hydrothermal systems or artificially creating a sufficient fracture system in a petrothermal reservoir may result in land subsidence or induce seismic events. Land subsidence was observed, e.g., at The Geysers, California, USA [205] or Wairakei, New Zealand [5, 6], whereas seismic events are recorded, e.g., for geothermal facilities in Landau, Germany [73], Basel, Switzerland [21], The Geysers [67, 125], Coso, California, USA [78], or Cerro Prieto, Mexico [114].

Focusing on high production rates, excessive withdrawal of energy can lead to depletion of a reservoir as happened at The Geysers [125], thus, risking the sustainability of the resource. This risk can be met by re-injection of withdrawn fluids, but may need a complex re-injection process increasing operational costs.

Operating a geothermal power plant, especially one based on steam-dominated reservoirs, requires sophisticated maintenance and operation management.

Flexible reactions to increases or decreases on demand of power are hard to manage and doing so may increase costs and reduce overall productivity. However, in a balanced renewable resources development strategy, such fluctuation may be addressed by other, complementing resources like wind or solar energy.

To meet the described challenges and reduce the exploratory as well as operational risks, cooperation between engineers and scientists from the disciplines of geology, (geo-)physics, geochemistry, (geo-)informatics, and geomathematics in a joint effort is needed [95]. In order to maximize success, profound knowledge of a potential reservoir, as, e.g., its geological, thermal, and mechanical properties, is equally important as the ability to predict reservoir behavior during drilling and operation. A condition for the latter is the development of a complex model (or set of several models) to describe the reservoir in mathematical terms. The Geomathematics Group at the University of Kaiserslautern developed a model to structure different tasks in the field of geothermal energy (Fig. 1.2) based on four columns: potential methods (gravity/geomagnetics) [84, 93, 104–108, 201], seismic exploration [83, 90, 134], transport processes [182, 214–216], and stress field [13, 14].

For an overview on the different columns of the model, see [15] and [17, 94] (in German). For further information on the mathematical background, see, e.g., [81, 82, 85, 86, 88, 91, 97] and the references therein.

This thesis is a contribution to the fourth column, the stress field simulations.

1.2 Literature on Rock Mechanics and Geomechanics

In order to describe a geothermal reservoir, the development of several quantities like (among others) temperature, (water) pressure, flow velocity, solid displacement, or density of solid and fluid parts has to be modeled. This may yield a complex set of (non-linear) partial differential equations, or even, depending on the effects that have to be incorporated, fractional, stochastic, or integro-differential equations. Such systems are often split into several subsystems and solved iteratively. As in the column model (Fig. 1.2), it is common practice to treat transport processes related to fluid or heat transport separately from processes involving deformation of the solid components. Such a splitting has to be incomplete as interactions between fluid pressure and rock stresses (poroelasticity) as well as temperature differences (thermoelasticity) influence the behavior of solid components. In this thesis, we concentrate on poroelastic effects in consolidation processes. Such effects are not only important during production phases of a geothermal facility, but already have to be considered for a correct determination of the stress field prior to drilling and production.

An introduction to rock mechanics is given, e.g., by the book of Jaeger, Cook, and Zimmerman [138], which also deals with fractures. General introductions to elasticity from a physical point of view can be found in, e.g., [169, 170], or from a mathematical point of view, e.g., in [188].

First attempts to incorporate the effects of fluids in a porous solid were made by von Terzaghi [271, 272], Frenkel [98], and Biot [34–38]. Since then, many different approaches were taken to develop a theory of poroelasticity that incorporates more general settings, e.g., even viscoelastic effects [282]. An overview and classification of theories in poroelasticity is given by Berryman in [32] who also identifies different ways of upscaling in order to get a macroscopic theory from microscopic considerations. Other recent considerations can be found in [39, 204, 260]. An elegant approach based on a Lagrangian formulation and representative elementary volume averaging can be found in [177–179]. The authors state that their final set of equations is equivalent to Biot's equations for homogeneous, consolidated (non-degenerated) materials with negligible porosity dynamics if at most linear waves are considered [178]. Consequently, the considerations in this thesis are based on Biot's equations as we are interested in consolidation phenomena.

Although some model problems in geomechanics can be solved analytically (see [251] and the references therein for some examples), most real-world applications have to be solved numerically. For an introduction to numerical methods in rock mechanics, see [139] and the references therein.

Overviews on rock mechanics issues with a focus on applications in geothermal systems can be found, e.g., in [71, 109]. The numerical methods used in the context of poroelasticity include finite element methods (FEM) [197, 209, 221–225], boundary element methods (BEM) [20, 51, 53, 54, 64, 110–112, 247, 290, 303], or hybrid Trefftz-finite-element-methods [202]. There are also approaches based on potential methods and Navier splines [87, 96] as well as Cauchy-Navier wavelets [1, 92] applied to deformation analysis, which, however, only treat problems of linear elasticity and still have to be generalized to poroelasticity. Models and numerical methods on rock failure and fracturing are presented, e.g., in [123, 130, 240, 304] and the references therein.

1.3 Outline

The aim of this thesis is the development of a numerical solution scheme with which fluid pressure and rock stresses in a geothermal reservoir can be determined prior to well drilling and during production. For this purpose, the method should

- Include poroelastic effects,
- Provide a possibility to include thermoelastic effects,
- Be not expensive in terms of memory and computational power,
- Be flexible with respect to location of data points.

The first point is met by taking Biot's equations of poroelasticity as point of departure. How to include thermoelastic effects into these is well known [286], but, as mentioned above, should be coupled with a model to describe heat and fluid transport. This coupling is out of the scope of this thesis and an interesting topic for further research. The third point in the list above can be satisfied by using a boundary method, preferably an integration-free one. The last point suggests the use of a meshless method, as it is hard to mesh complex geometries and as larger deformations make it necessary to remesh a domain. As a consequence, we do not consider traditional finite difference, finite element, or finite volume methods, as they all require a mesh. The same is true for the usual approach of boundary element methods, which may need even more sophisticated meshing algorithms.

In order to rise to the mentioned challenges, a fundamental solution scheme for Biot's equations of poroelasticity is developed. After proving that such a scheme can be used as a numerical solution method, we implement and test it for some examples. The content of the particular chapters of this thesis is as follows:

In Chap. 2, we start by summarizing the basic notation and mathematical results needed throughout this thesis, in particular from the disciplines of functional analysis, function spaces, linear operators, differential operators, integral calculus, and differential equations, including the Gauß (divergence) theorem, the theorems by Green, the Hahn-Banach theorem, and the Lax-Milgram theorem.

Next, Chap. 3 gives all the necessary basics from physics and mathematics to handle Biot's quasistatic equations of poroelasticity. Well known balance equations are combined with the constitutive relations suggested by Biot [34] to derive appropriate partial differential equations. This derivation is supplemented by some basic assumptions from linear elasticity to support a better understanding. By non-dimensionalizing Biot's equations, we show which terms may be neglected in the context of consolidation, and which must not. Based on the work by Showalter [254], we then take a look at the mathematical properties of Biot's quasistatic equations and derive an existence and uniqueness theorem with standard results from the theory of differential equations [175, 234]. In this derivation, we inspect a little bit further how some properties of non-homogeneous right-hand sides are transferred to properties of the solution. As a consequence, we derive a regularity result for the time derivative of the displacement vector, which could not yet be found in the literature.

Based on the previous results, Chap. 4 introduces boundary equations to the quasistatic equations of poroelasticity, similar to the well-known Green's formulas for, e.g., the Laplace equation [97], heat equation [57], Stokes equations [192], and Cauchy-Navier equation [91]. To this end, it is necessary to introduce the adjoint equations of quasistatic poroelasticity and fundamental solutions of the differential operator of quasistatic poroelasticity and its adjoint operator. This mathematically rigorous derivation is a novelty, as former derivations of boundary integral equations were based on a reciprocity relation [52]. Properties of the fundamental solutions are discussed and their relations to the aforementioned well-known differential equations are presented. Finally, it is shown how solutions of the homogeneous quasistatic equations of poroelasticity can be expressed by boundary integral equations and how equivalents to single- and double-layer potentials known from potential theory [97] can be deduced.

Solution schemes based on boundary integral formulations usually have to be regularized as these integrals are singular. Based on a regularization approach by Runge, Chap. 5 introduces the method of fundamental solutions by using the Laplace equation as a simple example. It is shown that this procedure can be readily generalized towards other equations, including the equations of poroelasticity, which allows us to introduce a method of fundamental solutions for them. Before doing this, we give a (necessarily incomplete) overview of some of the literature on the method of fundamental solutions, particularly with regard to density results and (rather scarce) convergence results. Based mainly on work by Smyrlis [262–264], we prove a density result for the method of fundamental solutions in poroelasticity developed from regularizing the corresponding boundary integral equations. To the best of our knowledge, density results for implicit evolution equations, systems of differential equations with a dependence on time, which are, thus, non-elliptic, or systems which cannot be easily decoupled into several separate equations have not been proven before.

By changing the variables in which Biot's equations are expressed, we will see that there is another way to derive a slightly different method for which density results are shown on the basis of results for the heat equation [143, 163] as well as the Cauchy-Navier equation [264] and explain how both fundamental solution methods are related. Although the result may look similar to considerations in [264], the difference is that, again, the system under consideration here contains an equation with a time-dependence, whereas in [264], only the static case is considered.

As an application, Chap. 6 presents some numerical results, starting with some background information on the implementation of the method of fundamental solutions in poroelasticity. Although the method of fundamental solutions was already applied to scalar-valued parabolic equations, explicitly the heat equation, and elliptic systems, we break ground with our attempt to apply this method for a system containing an implicit evolution equation. Next, we discuss the influence of different method parameters on the overall performance. For this purpose, four examples with different types of boundary conditions on the square $(-1, 1)^2$ in $\mathbb{R}^2$ are considered. The method parameters are organized in three groups. Each of these groups is discussed separately to determine which of them is the most critical with regard to a good approximation. Furthermore, a comparison is made between results obtained with a robust, but costly solution method for the system of linear equations which arises from our ansatz and a simple least-squares algorithm. As two slightly different approaches on a method of fundamental solutions are developed in Chap. 5, we investigate briefly which of those shows better numerical performance. Moreover, we will have a look at an alternative time-marching scheme and consider more challenging examples with steep gradients in order to test the limits of our method. Chapter 6 is rounded off by an exemplary application of the method of fundamental solutions for an initial boundary value problem on the three-dimensional cube $(-1, 1)^3$.

In conclusion, Chap. 7 summarizes the results of this thesis and gives an outlook on perspectives for future mathematical research in the context of poroelasticity and, in general, stress field modeling for geothermal reservoirs.

Chapter 2
Preliminaries

In this chapter, we will give a short overview on some basic concepts for the convenience of the reader. We start by introducing notation and summarize a few definitions as well as theorems from functional analysis, especially about linear operators, bilinear forms, and different kinds of function spaces. Moreover, several characterizations of the regularity of domains are given and a few theorems from vector analysis are recalled. The chapter concludes with some remarks on differential equations.

2.1 Basic Notation

As usual, we denote the set of positive integers by $\mathbb{N}$, the set of non-negative integers by $\mathbb{N}_0$, the set of all integers by $\mathbb{Z}$, the set of rational numbers by $\mathbb{Q}$, the set of real numbers by $\mathbb{R}$, positive real numbers by $\mathbb{R}^+$, non-negative real numbers by $\mathbb{R}_0^+$, and the set of complex numbers by $\mathbb{C}$. For any real number x, we define $\lfloor x \rfloor$ to be the largest integer n with $n \leq x$.

For any $n \in \mathbb{N}$, we define the n-dimensional real vector space as the cartesian product

$$\underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-times}} = \mathbb{R}^n \ .$$

For elements $x = (x_1, x_2, \ldots, x_n)^T, y = (y_1, y_2, \ldots, y_n)^T \in \mathbb{R}^n$, we define the (Euclidean) inner or scalar product $x \cdot y$ and its induced (Euclidean) norm $\|x\|$ to be

$$x \cdot y = \sum_{i=1}^{n} x_i y_i \ , \tag{2.1a}$$

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4_2

$$\|x\| = \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2} \ . \tag{2.1b}$$

Here, the upper index T means the transpose of a vector, and also denotes the transpose of a matrix later on. We frequently use Einstein's summation convention which means that using the same index variable twice within a single term implies summation of that term over all values of the index. For the Euclidean vector space $\mathbb{R}^3$, we also define the vector product $x \wedge y$ as

$$x \wedge y = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} \ . \tag{2.1c}$$

For any subset $\Omega \subset \mathbb{R}^n$, $\partial\Omega$ denotes its boundary and its closure is denoted by $\overline{\Omega} = \Omega \cup \partial\Omega$. Special kinds of subsets are balls. A ball of radius r centered at x will be denoted by $B_r(x)$ and is defined by

$$B_r(x) = \{y \in \mathbb{R}^n : \ \|x - y\| < r\} \ . \tag{2.2}$$

The corresponding sphere is given by $\partial B_r(x)$ or more explicitly

$$\partial B_r(x) = \{y \in \mathbb{R}^n : \ \|x - y\| = r\} \ . \tag{2.3}$$

We introduce two special indexed symbols. For $k, l \in \mathbb{Z}$, the Kronecker symbol δ_{kl} is defined by

$$\delta_{kl} = \begin{cases} 1, & k = l \ , \\ 0, & k \neq l \ . \end{cases} \tag{2.4}$$

The Levi-Civita tensor $\boldsymbol{\varepsilon}_{ijk}$, $i, j, k \in \{1, 2, 3\}$, is the total anti-symmetric tensor of order 3, given by

$$\boldsymbol{\varepsilon}_{ijk} = \begin{cases} +1, & \text{if } (i, j, k) \text{ is an even permutation of } (1, 2, 3) \ , \\ -1, & \text{if } (i, j, k) \text{ is an odd permutation of } (1, 2, 3) \ , \\ 0, & \text{otherwise} \end{cases} \tag{2.5a}$$

which yields

$$\boldsymbol{\varepsilon}_{ijk} = \frac{(i - j)(j - k)(k - i)}{2} \ . \tag{2.5b}$$

Here, we also introduced that we will use boldfaced letters for tensors of order 2 and higher.

2.2 Linear Operators and Bilinear Forms

Throughout this thesis, we will deal with a certain set of differential equations, i.e., differential operators and corresponding bilinear forms. Thus, in this section, we address linear operators and bilinear forms in a more abstract setting, starting with the definition of continuity and norms.

Definition 2.1 (Continuous Linear Operators and Operator Norm) Let $(V, \|\cdot\|_V)$ and $(Q, \|\cdot\|_Q)$ be two normed vector spaces. A linear operator $T: V \supset D(T) \to Q$ with domain $D(T) \subset V$ and range $R(T) \subset Q$ is continuous if and only if there exists a positive constant $C \in \mathbb{R}^+$ such that

$$\|Tx\|_Q \leq C \|x\|_V \text{ for all } x \in D(T) . \tag{2.6}$$

The space of all continuous linear operators from V to Q is denoted by $\mathscr{L}(V, Q)$. It is a normed space with norm given by

$$\|T\|_{\mathscr{L}(V,Q)} = \sup_{\|x\|_V=1} \|Tx\|_Q . \tag{2.7}$$

This norm is submultiplicative, i.e., for $T, S \in \mathscr{L}(V, Q)$, we have

$$\|TS\|_{\mathscr{L}(V,Q)} \leq \|T\|_{\mathscr{L}(V,Q)} \|S\|_{\mathscr{L}(V,Q)} . \tag{2.8}$$

If Q is a Banach space, then $\mathscr{L}(V, Q)$ is also a Banach space.

The kernel $\ker(T)$ and range $R(T)$ of a linear operator are defined as

$$\ker(T) = \{v \in V : Tv = 0\} , \tag{2.9}$$

$$R(T) = \{q \in Q : \exists v \in V : Tv = q\} . \tag{2.10}$$

Proof The submultiplicativity of the norm is proven, i.e., in [295, Chapter I.6].

Continuity of bilinear forms can be defined in a similar way.

Definition 2.2 (Continuous Bilinear Forms and their Norm) Let $(V, \|\cdot\|_V)$ and $(Q, \|\cdot\|_Q)$ be two normed vector spaces. A bilinear form $a(\cdot, \cdot) : V \times Q \to \mathbb{R}$ is continuous if and only if there exists a positive constant $C \in \mathbb{R}^+$ such that

$$|a(v, q)| \leq C \|v\|_V \|q\|_Q \text{ for all } v \in V, q \in Q . \tag{2.11}$$

The space of all continuous bilinear forms on $V \times Q$ is denoted by $\mathscr{L}(V \times Q, \mathbb{R})$ and is a normed space with a norm given by

$$\|a\|_{\mathscr{L}(V\times Q,\mathbb{R})} = \sup_{\|v\|_V=1} \sup_{\|q\|_Q=1} |a(v,q)| \ . \tag{2.12}$$

Proof The proof of submultiplicativity of the norm is done analogously to the previous one.

In order to establish a connection between bilinear forms and linear operators, we need to define the dual space V'.

Definition 2.3 (Dual Space and Dual Product) For a normed linear space $(V, \|\cdot\|_V)$, the set of all linear functionals on V is called its dual space and will be denoted by V'. The dual product of an element $v' \in V'$ and an element $v \in V$, denoted by

$$\left(v', v\right)_V = v'(v) \ , \tag{2.13}$$

is the value of the functional v' at the point v.

Lemma 2.4 (Linear Operator to a Bilinear Form) *For a continuous bilinear form $a(\cdot,\cdot) : V \times Q \to \mathbb{R}$ on the normed spaces $(V, \|\cdot\|_V)$ and $(Q, \|\cdot\|_Q)$ exists one and only one continuous linear operator $A : V \to Q'$ such that*

$$a(v,q) = (Av, q)_Q \tag{2.14}$$

for all $v \in V$ and $q \in Q$. The norms of A and $a(\cdot,\cdot)$ are equal.

Proof We can define $A : V \to Q'$ by imposing that for every $v \in V$ and $q \in Q$ the identity $(Av, q)_Q = a(v,q)$ is satisfied. Thus, A is well-defined. A is continuous and linear because $a(\cdot,\cdot)$ is continuous and (bi-)linear. Uniqueness of A results from (2.14).

In a similar way as the dual space allows us to define a linear operator corresponding to a bilinear form, it also allows us to define the adjoint operator of a linear operator.

Lemma 2.5 (Adjoint Operator) *Let $(V, \|\cdot\|_V)$ and $(Q, \|\cdot\|_Q)$ be two normed vector spaces. For any operator $T \in \mathscr{L}(V, Q)$, there exists a unique operator $T^* : Q' \to V'$ such that*

$$(T^* q', v)_V = (q', Tv)_Q \quad \textit{for all } v \in V \textit{ and } q' \in Q' \ . \tag{2.15}$$

The adjoint operator T^* is continuous, linear and its norm is the same as the norm of T.

If $(V, \|\cdot\|_V)$ and $(Q, \|\cdot\|_Q)$ are Banach spaces, the existence of an inverse operator $T^{-1} \in \mathscr{L}(Q, V)$ of T is equivalent to the existence of an inverse operator $(T^*)^{-1} \in \mathscr{L}(Q', V')$ of T^*. If the inverse operator exists, we have $(T^{-1})^* = (T^*)^{-1}$.

If T is only defined on a subset $D(T) \subsetneq V$, the above statements are still valid if T^* is defined on

$$D(T^*) = \{q' \in Q' : \ v \mapsto (q', Tv)_Q \text{ is continuous on } V\} \ . \tag{2.16}$$

Proof See, e.g., [295, Chapter VII].

It is sometimes necessary to not only consider linear operators, but families of linear operators, given by $(T(t))_{t\in\mathbb{R}_0^+}$. A useful restriction is to require such a family to be a strongly continuous semigroup [70, Definition 1.1]

Definition 2.6 (Strongly Continuous Semigroup) A family $(T(t))_{t\in\mathbb{R}_0^+}$ of bounded linear operators on a Banach space V is called a strongly continuous (one-parameter) semigroup if it satisfies

$$T(t+s) = T(t)T(s) \text{ for all } t, s \geq 0 \ , \tag{2.17}$$

$$T(0) = I \ , \tag{2.18}$$

with the identity operator I, and is strongly continuous, i.e., for every $v \in V$ the maps

$$t \mapsto T(t)v \tag{2.19}$$

are continuous from $\mathbb{R}^+$ to V for every $v \in V$.

In some cases, linear operators are not defined on a whole vector space V, but only on a subspace U. If the range of the operator is in $\mathbb{R}$, i.e., it is a linear functional, the Hahn-Banach Theorem gives a condition that an extension to the whole space exists.

Theorem 2.7 (Hahn-Banach Theorem) *Let U be a subspace of a real vector space V, $p : V \to \mathbb{R}$ with*

$$p(x+y) \leq p(x) + p(y) \ , \tag{2.20a}$$

$$p(tx) = tp(x) \tag{2.20b}$$

for all $x, y \in V$, $t \in \mathbb{R}_0^+$, and $f : \ U \to \mathbb{R}$ a linear functional with $f(x) \leq p(x)$ for all $x \in U$.
Then there exists a linear functional $F : V \to \mathbb{R}$ such that $F(x) = f(x)$ for all $x \in U$ and $-p(-x) \leq F(x) \leq p(x)$ for all $x \in V$.

If V is locally convex and f is continuous on U, then F is continuous on V.

Proof [See, e.g., 238, Theorems 3.2 and 3.6].

2.3 Function Spaces

In order to deal with differential equations, we have to introduce some notation for differentiation and integration. As it turns out, the classical strong concept of differentiability is too restrictive. This leads to the definition of weak differentiability. Different kinds of requirements on the differentiability of functions yield different sets of functions which can be shown to be normed vector spaces. It turns out that there are some more remarkable properties of these spaces and the functions they contain as well as interesting relations between them.

Definition 2.8 ((Strong) Differentiation) Let Ω be a bounded open subset of $\mathbb{R}^n$, $n \in \mathbb{N}$, $u : \Omega \to \mathbb{R}$, $\gamma \in \mathbb{N}_0^n$, and $k \in \mathbb{N}_0$. Let x be a point in $\mathbb{R}^n$ with coordinates x_i, $i \in \mathbb{N}$, $i \leq n$. Throughout this thesis, we will use cartesian coordinates.

The partial derivative $D^\gamma u$ is defined by

$$(D^\gamma u)(x) = \left(\partial_{x_1}^{\gamma_1} \dots \partial_{x_n}^{\gamma_n} u\right)(x) = \left(\frac{\partial^{|\gamma|} u}{\partial x_1^{\gamma_1} \dots \partial x_n^{\gamma_n}}\right)(x), \quad x \in \Omega \,, \tag{2.21}$$

with $|\gamma| = \sum_{i=1}^n \gamma_i$ being the order of the derivative. The set of all derivatives of u of order k at point x is denoted by $(D^k u)(x) = \{(D^\gamma u)(x) : |\gamma| = k\}$.

Remark 2.9

(i) Is u defined as a function on $\mathbb{R}^n$, $n \in \mathbb{N}$, such that $u : \mathbb{R}^n \to \mathbb{R}$, the restriction of u to $\Omega \subset \mathbb{R}^n$ is denoted by $u|_\Omega$.

(ii) The gradient of a differentiable function $u : \mathbb{R}^n \to \mathbb{R}$ is defined as the vector of all first derivatives and denoted by

$$\nabla_x u(x) = (\partial_{x_1} u, \dots, \partial_{x_n} u)^T \,. \tag{2.22}$$

We will omit the index x if it is clear with respect to which variable the differentiation has to be carried out.

(iii) The directional derivative of u with respect to a unit vector e is given by $(\nabla_x u) \cdot e$. The directional derivative with respect to the outer unit normal vector of a bounded domain Ω, i.e., a bounded connected open subset $\Omega \subset \mathbb{R}^n$ is denoted by $\partial_{n(x)} u$.

Please note the term "domain" as introduced here should not be confused with the domain of a linear operator as given in Definition 2.1.

As we now have introduced the notation for strong derivatives, we can define function spaces of continuously differentiable functions (see, e.g., [3]).

Definition 2.10 (Spaces of Continuously Differentiable Functions) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a domain, i.e. a connected open set. For $k \in \mathbb{N}_0$, we denote by $C^k(\Omega)$ the vector space of all functions $u : \Omega \to \mathbb{R}$ which together with all their derivatives $D^\gamma u$ of order $|\gamma| \leq k$ are continuous on Ω. Analogously, we define the space $\mathcal{C}^k(\Omega)$ containing all vector-valued functions $v = (v_1, v_2, v_3)^T : \Omega \to \mathbb{R}^3$ which together with all their derivatives $D^\gamma v_i$ of order $|\gamma| \leq k$ for $i \in \{1, 2, 3\}$ are continuous on Ω. We write $C(\Omega)$ for $C^0(\Omega)$ and $\mathcal{C}(\Omega)$ for $\mathcal{C}^0(\Omega)$.

The spaces of infinitely continuously differentiable functions are given by $C^\infty(\Omega) = \bigcap_{k=0}^\infty C^k(\Omega)$ and $\mathcal{C}^\infty(\Omega) = \bigcap_{k=0}^\infty \mathcal{C}^k(\Omega)$, respectively.

The subspace of functions in $C^k(\Omega)$ that have compact support in Ω is denoted by $C_0^k(\Omega)$ and analogously for vector-valued functions $\mathcal{C}_0^k(\Omega)$. A function u has compact support in Ω if there is a compact set $K \subset \Omega$ such that

$$\operatorname{supp}(u) = \overline{\{x \in \Omega : u(x) \neq 0\}} \subset K \,. \tag{2.23}$$

The spaces $C^k(\overline{\Omega})$ contain all functions $u \in C^k(\Omega)$ for which $D^\gamma u$ is bounded and uniformly continuous for all $\gamma \in \mathbb{N}_0^k$ with $0 \leq |\gamma| \leq k$. These spaces are Banach spaces when equipped with the norm

$$\|u\|_{C^k(\overline{\Omega})} = \max_{0 \leq |\gamma| \leq k} \sup_{x \in \Omega} |(D^\gamma u)(x)| \,. \tag{2.24}$$

The spaces $\mathcal{C}^k(\overline{\Omega})$ are defined accordingly.

Remark 2.11 Within this thesis, we reserve the term vector-valued for functions with values in $\mathbb{R}^2$ or $\mathbb{R}^3$.

In some cases, functions are required to be more regular than just being continuous, but requiring them to be continuously differentiable would be to much. Thus, we introduce the spaces of Hölder-continuous functions [3, 263].

Definition 2.12 (Hölder-Continuous Functions) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a domain, $\gamma \in \mathbb{N}_0^k$, and $k \in \mathbb{N}_0$. The space $C^{k,s}(\overline{\Omega})$, $0 < s \leq 1$, is the subspace of functions $u \in C^k(\overline{\Omega})$ whose derivatives $D^\gamma u$ of order k satisfy

$$|(D^\gamma u)(x) - (D^\gamma u)(y)| \leq C \|x - y\|^s \quad \forall x, y \in \Omega \tag{2.25}$$

with a constant $C \in \mathbb{R}^+$. We say u has Hölder-continuous derivatives of order k with Hölder exponent s or, in the special case $s = 1$, u has Lipschitz-continuous derivatives of order k. $C^{k,s}(\overline{\Omega})$ is a Banach space if equipped with the norm

$$\|u\|_{C^{k,s}(\overline{\Omega})} = \|u\|_{C^k(\overline{\Omega})} + \max_{0 \leq |\gamma| \leq k} \sup_{\substack{x,y \in \Omega \\ x \neq y}} \frac{|(D^\gamma u)(x) - (D^\gamma u)(y)|}{\|x - y\|^s} \,. \tag{2.26}$$

For $r \geq s$, the inclusion $C^{k,r}(\overline{\Omega}) \subset C^{k,s}(\overline{\Omega})$ is valid.

If a function u satisfies

$$\lim_{\delta \to 0} \sup_{\substack{x,y \in \Omega \\ 0 < \|x-y\| < \delta}} \frac{|(D^\gamma u)(x) - (D^\gamma u)(y)|}{\|x-y\|^s} \tag{2.27}$$

for each $|\gamma| = k$, the derivatives of order k of u are called uniformly Hölder-continuous and u is an element of the space $C_u^{k,s}(\overline{\Omega})$.

Proof See, e.g., [3, Theorem 1.31] for the inclusion.

When looking for functions whose values on the boundary of a domain are prescribed, it is often useful, if not even necessary, to restrict the kind of domain under consideration to answer questions of existence and uniqueness. The domain is required to have some kind of regularity. In order to give different kinds of regularity properties, we need another definition [3, Section 3.34].

Definition 2.13 (m-smooth Transformation) Let Φ be a one-to-one transformation of a domain $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, onto a domain $G \subset \mathbb{R}^n$ with $\Psi = \Phi^{-1}$. We call Φ m-smooth if, writing $y = \Phi(x)$ and

$$\begin{aligned} y_1 &= \phi_1(x_1, \ldots, x_n), & x_1 &= \psi_1(y_1, \ldots, y_n), \\ y_2 &= \phi_2(x_1, \ldots, x_n), & x_2 &= \psi_2(y_1, \ldots, y_n), \\ &\vdots & &\vdots \\ y_n &= \phi_n(x_1, \ldots, x_n), & x_n &= \psi_n(y_1, \ldots, y_n), \end{aligned}$$

the functions $\phi_1, \ldots, \phi_n$ belong to $C^m(\overline{\Omega})$ and the functions $\psi_1, \ldots, \psi_n$ belong to $C^m(\overline{G})$.

The following summary of regularity conditions is taken from the book by Adams [3, pp. 66f] and is just slightly adapted. Here, a finite cone with vertex x is defined as the set

$$\mathscr{C}_x = B_{r_1}(x) \cap \{x + \lambda(y-x) : \ y \in B_{r_2}(z), \lambda > 0\} \tag{2.28}$$

with $B_{r_2}(z)$ being a ball such that $x \notin B_{r_2}(z)$ and an open cover $\{U_j\}$ of a set Ω is said to be locally finite if any compact set in $\mathbb{R}^n$ can intersect at most finitely many elements of $\{U_j\}$[3, p. 65ff].

Definition 2.14 (Regularity of Domains) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a domain. Ω has

(i) The *segment property* if there exists a locally finite open cover $\{U_j\}$ of $\partial\Omega$ and a corresponding sequence $\{y_j\}$ of non-zero vectors such that if $x \in \overline{\Omega} \cap U_j$ for some j, then $x + \varepsilon y_j \in \Omega$ for $0 < \varepsilon < 1$,

(ii) The *cone property* if there exists a finite cone $\mathscr{C}$ such that each point $x \in \Omega$ is the vertex of a finite cone $\mathscr{C}_x$ contained in Ω and congruent to $\mathscr{C}$.

(iii) The *uniform cone property* if there exists a locally finite open cover $\{U_j\}$ of $\partial\Omega$ and a corresponding sequence $\{\mathscr{C}_j\}$ of finite cones, each congruent to some fixed finite cone $\mathscr{C}$, such that

  (a) For some finite $M \in \mathbb{R}^+$, every U_j has a diameter less than M,
  (b) For some $\delta > 0$, $\bigcup_{j=1}^{\infty} U_j \supset \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) < \delta\}$,
  (c) For every j, $Q_j = \bigcup_{x \in \Omega \cap U_j} (x + \mathscr{C}_j) \subset \Omega$,
  (d) For some finite $R \in \mathbb{N}$, every collection of $R + 1$ of the sets Q_j has an empty intersection.

(iv) The *strong local Lipschitz property* if there exist positive numbers δ and M, a locally finite open cover $\{U_j\}$ of $\partial\Omega$, and for each U_j a real-valued function f_j of $n - 1$ real variables, such that

  (a) For some finite $R \in \mathbb{N}$, every collection of $R + 1$ of the sets U_j has an empty intersection,
  (b) For every pair of points $x, y \in \{z \in \Omega : \operatorname{dist}(z, \partial\Omega) < \delta\}$ such that $\|x - y\| < \delta$, there exists j such that $x, y \in \{z \in U_j : \operatorname{dist}(z, \partial U_j) > \delta\}$,
  (c) Each function f_j satisfies a Lipschitz condition with constant M,
  (d) For some cartesian coordinate system $(\xi_{j,l})_{l=1}^n$ in U_j, the set $\Omega \cap U_j$ is represented by the inequality $\xi_{j,n} < f_j(\xi_{j,1}, \ldots, \xi_{j,n-1})$.

(v) The *uniform C^m-regularity property* if there exists a locally finite open cover $\{U_j\}$ of $\partial\Omega$ and a corresponding sequence $\{\Phi_j\}$ of m-smooth one-to-one transformations with Φ_j taking U_j onto $B_1(0) \subset \mathbb{R}^n$, such that

  (a) For some $\delta > 0$, $\bigcup_{j=1}^{\infty} \Psi_j(B_{0.5}(0)) \supset \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) < \delta\}$, where $\Psi = \Phi^{-1}$,
  (b) For some finite $R \in \mathbb{N}$, every collection of $R + 1$ of the sets U_j has an empty intersection,
  (c) For each j, $\Phi_j(U_j \cap \Omega) = \{y \in B_1(0) : y_n > 0\}$,
  (d) If $(\phi_{j,1}, \ldots, \phi_{j,n})$ and $(\psi_{j,1}, \ldots, \psi_{j,n})$ denote the components of Φ_j and Ψ_j, respectively, then there exists a finite M such that for all $\gamma \in \mathbb{N}_0^n$, $|\gamma| \le m$, for every $1 \le i \le n$, and for every j, we have $\left|D^\gamma \phi_{j,i}(x)\right| \le M$, $x \in U_j$, and $\left|D^\gamma \psi_{j,i}(y)\right| \le M$, $y \in B_1(0)$.

For the different kinds of regularity, we have (v) $\overset{m \ge 1}{\Longrightarrow}$ (iv) $\Longrightarrow$ (iii) $\Longrightarrow$ (i). These regularity properties require Ω to lie on only one side of its boundary, whereas the cone property does not impose this condition.

Remark 2.15

(i) If Ω is bounded, the requirements for Ω being strong local Lipschitz reduce to the condition that for each point $x \in \partial\Omega$, there exists a neighborhood U of x such that $U \cap \partial\Omega$ is the graph of a Lipschitz-continuous function.

(ii) In some cases it is necessary to require that the parts of the one-to-one transformation mentioned in the definition of the C^m-regularity property have not only bounded derivatives, but Hölder-continuous ones. This yields the $C^{m,s}$-regularity property.

As already mentioned, the above introduced definition of strong differentiability with continuous or even Hölder-continuous derivatives is often too restrictive. Therefore, we need some other, weaker definition of derivatives. To define these weak derivatives, we need a definition of convergence in $C_0^\infty(\Omega)$ first (see, e.g., [3, Section 1.51]).

Definition 2.16 (Convergence in $C_0^\infty(\Omega)$) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a bounded domain, $\{\phi_l\}_{l\in\mathbb{N}_0} \subset C_0^\infty(\Omega)$, and $\phi \in C_0^\infty(\Omega)$. The sequence $\{\phi_l\}_{l\in\mathbb{N}_0}$ is said to converge towards ϕ in $C_0^\infty(\Omega)$ for $l \to \infty$ if there is a compact subset $K \subset \Omega$ such that

$$\operatorname{supp}(\phi_l) \subset K \text{ for all } l \in \mathbb{N}_0 \,, \tag{2.29a}$$

$$\operatorname{supp}(\phi) \subset K \,, \tag{2.29b}$$

as well as all partial derivatives of ϕ_l of arbitrary order converge uniformly to those of ϕ, i.e.,

$$\sup_{x\in\Omega} |(D^\gamma \phi_l)(x) - D^\gamma \phi(x)| \xrightarrow{l\to\infty} 0 \text{ for all } \gamma \in \mathbb{N}_0^n \,. \tag{2.29c}$$

The equivalent definition holds in $\mathcal{C}_0^\infty(\Omega)$.

Remark 2.17 $C_0^\infty(\Omega)$ is often denoted by $\mathscr{D}(\Omega)$ and called the space of test functions, although the latter identification is not unique. It is a topological vector space, or to be more precise, a Fréchet space, but not normable [3, 238].

The above definition allows us to define distributions (see [234, Definition 5.8]).

Definition 2.18 (Distribution) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a bounded domain. A distribution (or generalized function) is a linear functional $f : C_0^\infty(\Omega) \to \mathbb{R}$ which is continuous in the following sense: If a sequence $\{\phi_l\}_{l\in\mathbb{N}_0} \subset C_0^\infty(\Omega)$ converges for $l \to \infty$ towards $\phi \in C_0^\infty(\Omega)$, then $f(\phi_l) = (f, \phi_l)_{C_0^\infty(\Omega)}$ converges for $l \to \infty$ towards $f(\phi)$.

The set of all distributions is denoted by $\left(C_0^\infty(\Omega)\right)'$.

The space of vector-valued distributions is defined accordingly.

Remark 2.19 The space of distributions is often denoted by $\mathscr{D}'(\Omega)$. If we consider $C_0^\infty(\Omega)$ as a topological vector space, $\left(C_0^\infty(\Omega)\right)'$ is its topological dual, i.e., it consists of all linear functionals on $C_0^\infty(\Omega)$ which are continuous with respect to the weak topology [3, 238].

There is another way to characterize functions that is useful to present here in anticipation of a more general concept that we will introduce later on. For this, we have to explain our interpretation of the integral of a function.

Within this thesis, all integrals are understood in the sense of Lebesgue integrals. Note that the choice of the appropriate Lebesgue measure depends on the domain of integration. We mention this at this point because we will later on encounter integrals over the boundary of three-dimensional domains which would, of course, vanish if integration would be performed with respect to the Lebesgue measure in $\mathbb{R}^3$.

Now, we can define [3, Section 1.53].

Definition 2.20 (Locally Integrable Functions) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a bounded domain. A function u is called locally integrable on Ω if for every compact subset $K \subset \Omega$ we have

$$\int_K |f(x)| \, \mathrm{d}V_n(x) < \infty \,, \tag{2.30}$$

where $\mathrm{d}V_n$ denotes the volume element in $\mathbb{R}^n$.

Remark 2.21 If there is no confusion, we will omit the index n in $\mathrm{d}V_n$.

For every locally integrable function u, we can define a corresponding distribution $T_u \in \left(C_0^\infty(\Omega)\right)'$ simply by

$$T_u(\phi) = \int_\Omega u(x)\phi(x) \, \mathrm{d}V(x), \quad \phi \in C_0^\infty(\Omega) \,. \tag{2.31}$$

Usually, notation is a little bit abused by also using u instead of T_u to denote the corresponding distribution.

The reverse of the last statement is not true. There are many distributions for which no corresponding locally integrable function can be found. The most prominent example is the evaluation of a function ϕ at a certain point x, known as Dirac's delta distribution. If $0 \in \Omega$, the evaluation of a function $\phi \in C_0^\infty(\Omega)$ is given by

$$\delta(\phi) = \phi(0). \tag{2.32}$$

It is easy to prove that there is no locally integrable function for which

$$\int_\Omega \delta(x)\phi(x) \, \mathrm{d}V(x) = \phi(0), \quad \phi \in C_0^\infty(\Omega) \,. \tag{2.33}$$

However, δ obviously satisfies Definition 2.18.

In the course of this thesis, especially in Chap. 5, we need some operations on distributions. It is obvious how addition of two distributions and multiplication with a constant should be defined on distributions. Distributions may even be multiplied

by smooth functions [3, Section 1.58]. For $T \in \left(C_0^\infty(\Omega)\right)'$ and $u \in C_0^\infty(\Omega)$, the product $uT \in \left(C_0^\infty(\Omega)\right)'$ is defined by

$$(uT)(\phi) = T(u\phi), \quad \phi \in C_0^\infty(\Omega) \,. \tag{2.34}$$

The support of a distribution is defined as follows [238, Definition 6.22].

Definition 2.22 (Support of a Distribution) Suppose $T \in \left(C_0^\infty(\Omega)\right)'$ for an open bounded domain $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$. If ω is a subset of Ω and if $T\phi = 0$ for all $\phi \in C_0^\infty(\omega)$, we say that T vanishes on ω. Let W be the union of all sets $\omega \subset \Omega$ on which T vanishes. The complement of W relative to Ω is the support of T.

Obviously, the definition of compact support now also holds for distributions.

Another operation which we need not only on distributions is convolution [238, Definition 6.25].

Definition 2.23 (Convolution) Let u be a function defined on $\mathbb{R}^n$, $n \in \mathbb{N}$, and $x \in \mathbb{R}^n$. We define

$$(\tau_x u)(y) = u(y - x), \quad y \in \mathbb{R}^n \,, \tag{2.35}$$

$$\check{u}(y) = u(-y) \,, \tag{2.36}$$

$$(\tau_x \check{u})(y) = u(x - y) \,. \tag{2.37}$$

Let v be another function on $\mathbb{R}^n$. The convolution $u * v$ is defined as

$$(u * v)(x) = \int_{\mathbb{R}^n} u(y)v(x - y)\, \mathrm{d}V(y) = \int_{\mathbb{R}^n} u(y)(\tau_x \check{v})(y)\, \mathrm{d}V(y) \tag{2.38}$$

if the integral exists for almost all $x \in \mathbb{R}^n$.

For a distribution $u \in \left(C_0^\infty(\mathbb{R}^n)\right)'$ and $\phi \in C_0^\infty(\mathbb{R}^n)$, the function $u * \phi$ is given by

$$(u * \phi)(x) = u(\tau_x \check{\phi}). \tag{2.39}$$

Remark 2.24 If we have $u \in C_0^\infty(\mathbb{R}^3 \times \mathbb{R}_0^+)$, we define $\check{u}(x, t) = u(-x, -t)$.

Theorem 2.25 (Properties of Convolutions) *Let $u \in \left(C_0^\infty(\mathbb{R}^n)\right)'$, $\phi \in C_0^\infty(\mathbb{R}^n)$, $\psi \in C_0^\infty(\mathbb{R}^n)$. Then*

(i) $\tau_x(u * \phi) = (\tau_x u) * \phi = u * (\tau_x \phi)$ *for all* $x \in \mathbb{R}^n$,
(ii) $u * \phi \in C^\infty(\mathbb{R}^n)$ *and* $u * (\phi * \psi) = (u * \phi) * \psi$.
(iii) The operator L, defined by

$$L\phi = u * \phi, \quad \phi \in C_0^\infty(\mathbb{R}^n) \,, \tag{2.40}$$

is a continuous linear mapping of $C_0^\infty(\mathbb{R}^n)$ *into* $C^\infty(\mathbb{R}^n)$ *which satisfies* $\tau_x L = L\tau_x$, $x \in \mathbb{R}^n$.

(iv) If L is a continuous linear mapping of $C_0^\infty(\mathbb{R}^n)$ into $C(\mathbb{R}^n)$ which satisfies $\mathsf{T}_x L = L\mathsf{T}_x$, $x \in \mathbb{R}^n$, then there exists a unique $u \in \left(C_0^\infty(\mathbb{R}^n)\right)'$ such that L has a representation like (2.40).

Proof See [238, Theorems 6.30, 6.33].

Until now, convolutions for distributions are only declared if a distribution is convolved with an element of $C_0^\infty(\mathbb{R}^n)$. The next lemma extends convolution to elements of $C^\infty(\mathbb{R}^n)$.

Lemma 2.26 *Let $u \in \left(C_0^\infty(\mathbb{R}^n)\right)'$ have compact support, $\phi \in C^\infty(\mathbb{R}^n)$, $\psi \in C_0^\infty(\mathbb{R}^n)$. The convolution $u * \phi \in C^\infty(\mathbb{R}^n)$ is well-defined. Moreover,*

(i) $\mathsf{T}_x(u * \phi) = (\mathsf{T}_x u) * \phi = u * (\mathsf{T}_x \phi)$ *for all* $x \in \mathbb{R}^n$,
(ii) $u * (\phi * \psi) = (u * \phi) * \psi$,
(iii) $u * \psi \in C_0^\infty(\mathbb{R}^n)$,
(iv) $u * (\phi * \psi) = (u * \phi) * \psi = (u * \psi) * \phi$.

Proof See [238, Theorem 6.35].

Convolutions may also be defined between distributions.

Lemma 2.27 (Convolutions between Distributions) *Let $u, v, w \in \left(C_0^\infty(\mathbb{R}^n)\right)'$, $n \in \mathbb{N}$.*

(i) If at least one of u, v has compact support, the convolution $u * v$ is defined by $(u * v) * \phi = u * (v * \phi)$ for all $\phi \in C_0^\infty(\mathbb{R}^n)$ and $u * v = v * u$.
(ii) If at least one of the supports $\operatorname{supp}(u)$, $\operatorname{supp}(v)$ *is compact, we have* $\operatorname{supp}(u * v) \subset \operatorname{supp}(u) + \operatorname{supp}(v)$.
(iii) If at least two of the supports $\operatorname{supp}(u)$, $\operatorname{supp}(v)$, $\operatorname{supp}(w)$ *are compact, we have $(u * v) * w = u * (v * w)$.*

Proof See [238, Definition 6.36, Theorem 6.37].

We can now define weak derivatives by defining derivatives of distributions [3, Section 1.57], [238, Section 6.12].

Definition 2.28 (Weak Derivative) Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $u \in \left(C_0^\infty(\Omega)\right)'$. The weak derivative of u with respect to x_i, $i \in \{1, \ldots, n\}$, is defined by

$$(\partial_{x_i} u, \phi)_{C_0^\infty(\Omega)} = -(u, \partial_{x_i}\phi)_{C_0^\infty(\Omega)}, \quad \phi \in C_0^\infty(\Omega). \tag{2.41}$$

For a multi-index $\gamma \in \mathbb{N}_0^n$, we have the generalization

$$(D^\gamma u, \phi)_{C_0^\infty(\Omega)} = (-1)^{|\gamma|} (u, D^\gamma \phi)_{C_0^\infty(\Omega)}, \quad \phi \in C_0^\infty(\Omega). \tag{2.42}$$

Remark 2.29 We use the same notation for weak derivatives and classical (strong) partial derivatives (based on the limit of difference quotients). If a continuous strong

derivative exists, it coincides with the weak derivative as can be seen by integration by parts.

Theorem 2.30 (Weak Derivatives and Convolution)

(i) Suppose $u \in \left(C_0^\infty(\mathbb{R}^n)\right)'$ *and* $\phi \in C_0^\infty(\mathbb{R}^n)$, $n \in \mathbb{N}$, *or* $u \in \left(C_0^\infty(\mathbb{R}^n)\right)'$ *with compact support and* $\phi \in C^\infty(\mathbb{R}^n)$, *then* $D^\gamma(u * \phi) = (D^\gamma u) * \phi = u * (D^\gamma \phi)$ *for all* $\gamma \in \mathbb{N}_0^n$.

(ii) Suppose $u \in \left(C_0^\infty(\mathbb{R}^n)\right)'$ *and* δ *is the delta distribution, then* $D^\gamma u = (D^\gamma \delta)u$ *for all* $\gamma \in \mathbb{N}_0^n$. *In particular,* $u = \delta * u$.

(iii) Suppose $u, v \in \left(C_0^\infty(\mathbb{R}^n)\right)'$ *and at least one of them has compact support, then* $D^\gamma(u * v) = (D^\gamma u) * v = u * (D^\gamma v)$ *for all* $\gamma \in \mathbb{N}_0^n$.

Proof See [238, Theorems 6.30, 6.35, and 6.37].

A consequence of Theorem 2.30(ii) is that it allows to give an integral expression for Dirac's delta distribution and its derivatives, with a slight abuse of notation, by

$$\int_{\mathbb{R}^n} D^\gamma \delta(x-y)\phi(y)\,\mathrm{d}V(y) = (-1)^{|\gamma|}(D^\gamma \phi)(x)\ . \tag{2.43}$$

It is easy to proof that the weak derivative of a distribution is also a distribution. Thus, for every distribution, there exist weak derivatives of arbitrary order. Nevertheless, classes of distributions and their derivatives can be distinguished if we introduce a new concept of regularity based on integrability. We begin by defining Lebesgue spaces.

Definition 2.31 (Lebesgue Spaces) Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $p \in \mathbb{R}^+$. The Lebesgue space $L^p(\Omega)$ consists of all equivalence classes with respect to the Lebesgue measure of almost everywhere identical functions on Ω, whose representatives u satisfy

$$\int_\Omega |u(x)|^p\ \mathrm{d}V(x) < \infty\ . \tag{2.44}$$

Moreover, the space $L^\infty(\Omega)$ contains all such equivalence classes whose representatives are measurable, essentially bounded functions $u : \Omega \to \mathbb{R}$ with

$$\operatorname*{ess\,sup}_{x\in\Omega} |u(x)| < \infty\ . \tag{2.45}$$

Equivalent definitions hold for the spaces $\mathcal{L}^p(\Omega)$ and $\mathcal{L}^\infty(\Omega)$ of vector-valued functions.

Remark 2.32 It is convenient to identify a function with its respective equivalence class.

We summarize a few properties of the Lebesgue spaces.

Lemma 2.33 (Properties of Lebesgue Spaces) *Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $1 \leq p < \infty$.*

(i) The Lebesgue space $L^p(\Omega)$ is a Banach space with respect to the norm

$$\|u\|_{L^p(\Omega)} = \left(\int_\Omega |u(x)|^p \, \mathrm{d}V(x) \right)^{\frac{1}{p}} . \tag{2.46}$$

$L^\infty(\Omega)$ is a Banach space with respect to the norm

$$\|u\|_{L^\infty(\Omega)} = \operatorname*{ess\,sup}_{x \in \Omega} |u(x)| . \tag{2.47}$$

(ii) The space $L^2(\Omega)$ is a Hilbert space if equipped with the scalar product

$$(u, v)_{L^2(\Omega)} = \int_\Omega u(x) v(x) \, \mathrm{d}V(x) . \tag{2.48}$$

(iii) For arbitrary $1 \leq p_1, p_2 < \infty$ with $p_1 \geq p_2$, we have $L^{p_1}(\Omega) \subset L^{p_2}(\Omega)$ and $L^{p_1}(\Omega) \subset L^1_{\mathrm{loc}}(\Omega)$, with $L^1_{\mathrm{loc}}(\Omega)$ being the space of locally integrable functions.

(iv) Let $1 < p_1 < \infty$ and p_2 such that $\frac{1}{p_1} + \frac{1}{p_2} = 1$. For $u \in L^{p_1}(\Omega)$, $v \in L^{p_2}(\Omega)$, we have $uv \in L^1(\Omega)$ and

$$\|uv\|_{L^1(\Omega)} \leq \|u\|_{L^{p_1}(\Omega)} \|v\|_{L^{p_2}(\Omega)} . \tag{2.49}$$

This is known as Hölder's inequality. It also holds for $u \in L^1(\Omega)$ and $v \in L^\infty(\Omega)$; we have $uv \in L^1(\Omega)$.

(v) Let $1 < p_1 < \infty$ and p_2 such that $\frac{1}{p_1} + \frac{1}{p_2} = 1$. The dual spaces of Lebesgue spaces are given by

$$(L^{p_1}(\Omega))' = L^{p_2}(\Omega) . \tag{2.50}$$

Moreover, $\left(L^1(\Omega)\right)' = L^\infty(\Omega)$, but $(L^\infty(\Omega))' \neq L^1(\Omega)$.

(vi) $C_0(\Omega)$ and $C_0^\infty(\Omega)$ are dense subspaces of $L^p(\Omega)$ for all $1 \leq p \leq \infty$.

All of the above statements are also valid for spaces $\mathcal{L}^p(\Omega)$ with the obvious modifications.

Proof See, e.g., [3, Chapter 2].

The definition of Lebesgue spaces allows us to evaluate the regularity of a distribution by asking if it is also an element of some Lebesgue space. It comes naturally to extend this to a distribution's derivative. This gives rise to the definition of Sobolev spaces [194, 292].

Definition 2.34 (Sobolev Spaces) Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $1 \leq p \leq \infty$. The Sobolev space $W^{k,p}(\Omega)$, $k \in \mathbb{N}_0$, is defined as the subspace of $L^p(\Omega)$ with

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) : \ D^\gamma u \in L^p(\Omega) \text{ for all } \gamma \in \mathbb{N}_0^n, |\gamma| \leq k\} \ . \tag{2.51}$$

$W^{k,p}(\Omega)$ is a separable Banach space with respect to the norm

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\gamma| \leq k} \int_\Omega |D^\gamma u(x)|^p \ \mathrm{d}V(x) \right)^{\frac{1}{p}} \ . \tag{2.52}$$

For $p = 2$, we denote $H^k(\Omega) = W^{k,2}(\Omega)$. These spaces are separable Hilbert spaces with inner product

$$(u, v)_{H^k(\Omega)} = \sum_{|\gamma| \leq k} (D^\gamma u, D^\gamma v)_{L^2(\Omega)} \ . \tag{2.53}$$

Sobolev spaces of vector-valued functions are denoted by $\mathcal{W}^{k,p}(\Omega)$ as well as $\mathcal{H}^p(\Omega)$ and are defined analogously.

Proof Proofs of Sobolev spaces being Banach spaces or even Hilbert spaces can be found, e.g., in [3, Chapter III] or [292, Chapter 3].

Moreover, the concept of Hölder-continuity can also be transferred to weak derivatives in the following sense.

Definition 2.35 (Sobolev-Slobodeckij Spaces) Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $1 \leq p \leq \infty$. The Sobolev-Slobodeckij space $W^{k,p}(\Omega)$ of fractional order $k = r + s$ with $r \in \mathbb{N}_0$ and $0 < s < 1$ is defined as the subspace of $W^{r,p}(\Omega)$ with

$$W^{k,p}(\Omega) = \left\{u \in W^{r,p}(\Omega) : \ |D^\gamma u|_{s,p,\Omega} < \infty \text{ for all } \gamma \in \mathbb{N}_0^n, |\gamma| = k\right\} \ , \tag{2.54}$$

where the semi-norm $|u|_{s,p,\Omega}$ is given by

$$|u|_{s,p,\Omega} = \left(\int_\Omega \int_\Omega \frac{|u(x) - u(y)|^p}{\|x - y\|^{n+ps}} \ \mathrm{d}V(x) \ \mathrm{d}V(y) \right)^{\frac{1}{p}} \ . \tag{2.55}$$

$W^{k,p}(\Omega)$ is a Banach space if equipped with the norm

$$\|u\|_{W^{k,p}(\Omega)} = \left(\|u\|^p_{W^{r,p}(\Omega)} + \sum_{|\gamma| = k} |D^\gamma u|^p_{s,p,\Omega} \right)^{\frac{1}{p}} \ . \tag{2.56}$$

$\mathcal{W}^{k,p}(\Omega)$ for vector-valued functions is defined accordingly.

Proof For the proof of completeness, see, e.g., [292, Chapter 3].

In what follows, we formulate results only for scalar-valued functions, although similar results are valid for vector-valued functions.

The following relations of Sobolev spaces and spaces of continuous differentiable functions are particularly useful when considering numerical solution schemes.

Lemma 2.36 *Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $1 \leq p < \infty$. The Sobolev space $W^{k,p}(\Omega)$, $k \in \mathbb{N}_0$, is the completion of $C^\infty(\Omega)$ with respect to the norm $\|\cdot\|_{W^{k,p}(\Omega)}$.*

If Ω has the segment property, then the set of restrictions to Ω of functions in $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\Omega)$.

Proof See [3, Theorems 3.16, 3.18].

Lemma 2.36 suggests the definition of another class of Sobolev spaces.

Definition 2.37 Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $1 \leq p < \infty$. The Sobolev space $W_0^{k,p}(\Omega)$, $k \in \mathbb{N}_0$, is defined as the completion of $C_0^\infty(\Omega)$ with respect to the norm $\|\cdot\|_{W^{k,p}(\Omega)}$.

Remark 2.38 In general, $W_0^{k,p}(\Omega) \neq W^{k,p}(\Omega)$. For conditions on Ω under which those spaces are equal, the reader is referred to, e.g., [3, p. 56ff].

With this definition, we can characterize Sobolev spaces with negative index.

Definition 2.39 (Sobolev Spaces with Negative Index) Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $1 < p_1 < \infty$, p_2 such that $\frac{1}{p_1} + \frac{1}{p_2} = 1$. The Sobolev space $W^{-k,p_2}(\Omega)$, $k \in \mathbb{R}^+$, is defined as

$$W^{-k,p_2}(\Omega) = \left\{ f \in \left(C_0^\infty(\Omega)\right)' : \ \|f\|_{W^{-k,p_2}(\Omega)} < \infty \right\} , \tag{2.57}$$

with

$$\|f\|_{W^{-k,p_2}(\Omega)} = \sup_{0 \neq u \in C_0^\infty(\Omega)} \frac{|f(u)|}{\|u\|_{W^{k,p_1}(\Omega)}} . \tag{2.58}$$

$W^{-k,p_2}(\Omega)$ is the dual space of $W^{k,p_1}(\Omega)$.

Proof For norm properties and the duality relation, see, e.g., [3, 3.10–3.13].

An essential property of Sobolev spaces is the existence of the following imbeddings:

Theorem 2.40 (Sobolev Imbedding Theorem) *Let Ω be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}$, $j, k \in \mathbb{N}_0$, $1 \leq p_1, p_2 < \infty$.*

(i) If Ω has the cone property, the following imbeddings exist:

(a) Suppose $kp_1 < n$ and $p_1 \leq p_2 \leq \frac{np_1}{n-kp_1}$. Then

$$W^{j+k,p_1}(\Omega) \hookrightarrow W^{j,p_2}(\Omega) . \tag{2.59}$$

(b) Suppose $kp_1 = n$, $p_1 \le p_2 < \infty$. Then

$$W^{j+k,p_1}(\Omega) \hookrightarrow W^{j,p_2}(\Omega) \,. \tag{2.60}$$

Moreover, if $p_1 = 1$ and, thus, $k = n$, this also holds for $p_2 = \infty$.

(ii) If Ω has the strong local Lipschitz property, additional imbeddings hold:

(a) Suppose $kp_1 > n > (k-1)p_1$. Then

$$W^{j+k,p_1}(\Omega) \hookrightarrow C^{j,s}(\overline{\Omega}), \quad 0 < s < k - \frac{n}{p_1} \,. \tag{2.61}$$

(b) Suppose $n = (k-1)p_1$. Then

$$W^{j+k,p_1}(\Omega) \hookrightarrow C^{j,s}(\overline{\Omega}), \quad 0 < s < 1 \,. \tag{2.62}$$

The last imbedding holds for $s = 1$ if $n = k - 1$ and $p_1 = 1$.

Proof See, e.g., [3, Chapter V].

We will later on encounter the necessity to specify in some sense the values an element of a Sobolev space takes on the boundary of a domain Ω. This is not a trivial problem as elements of Sobolev spaces are equivalence classes like the elements of Lebesgue spaces on which the definition of Sobolev spaces is based. Lemma 2.36 allows us to find a solution to this dilemma by introducing the trace operator.

Theorem 2.41 (Trace Operator) *Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a bounded domain with the uniform $C^{m,s}$-regularity property.*

(i) Let $\frac{1}{2} < k \le m + s$, whereas for $k \in \mathbb{N}$, $k = m - 1$, $s = 1$ is allowed. There is a continuous linear operator $T_0 : H^k(\Omega) \to H^{k-\frac{1}{2}}(\partial\Omega)$, called trace operator, such that

$$T_0 u = u|_{\partial\Omega} \text{ for all } u \in C^{\lfloor k \rfloor + 1}(\overline{\Omega}) \,. \tag{2.63}$$

If $k \in \mathbb{N}$, we have $u \in C^k(\overline{\Omega})$.

(ii) Let $k + 1 \le m + s$, whereas for $k \in \mathbb{N}$, $m = k$ and $s = 1$ is allowed and $l \in \mathbb{N}$ such that $k - l > \frac{1}{2}$. There is another continuous linear trace operator $T_l : H^k(\Omega) \to \bigtimes_{i=0}^{l} H^{k-i-\frac{1}{2}}(\partial\Omega)$ such that

$$T_l u = \left(u|_{\partial\Omega}, \partial_{-n(x)} u|_{\partial\Omega}, \ldots, \partial^l_{-n(x)} u|_{\partial\Omega} \right) \text{ for all } u \in C^{\lfloor k \rfloor + l + 1}(\overline{\Omega}) \,. \tag{2.64}$$

If $k \in \mathbb{N}$, we have $u \in C^{k+l}(\overline{\Omega})$. Here, $\partial_{-n(x)} u$ is the directional derivative with respect to the inner normal on $\partial\Omega$.

Proof See [292, Theorem 8.7], [175, Chapter I, Theorem 3.2] and the references therein.

Remark 2.42 Definition 2.37 and Theorem 2.41 are compatible, as we have

$$T_0 u = 0 \quad \text{on } \partial\Omega \ \forall u \in W_0^{k,p}(\Omega) \ . \tag{2.65}$$

For $k \geq 2$, it is possible to show that

$$T_0 D^\gamma u = 0 \quad \text{on } \partial\Omega \ \forall u \in W_0^{k,p}(\Omega) \text{ with } |\gamma| \leq k-1 \ . \tag{2.66}$$

The definition of all the above spaces can be generalized to functions which take values in a Banach space X [234, Chapter 10]. Let $I \subset \mathbb{R}$ be a bounded open interval and X be a Banach space with (topological) dual X'. We start by defining $C(I;X)$ to be the space of all bounded continuous functions $u : I \to X$, $t \mapsto u(t)$ and equip it with the norm

$$\|u\|_{C(I;X)} = \sup_{t \in I} \|u(t)\|_X \ . \tag{2.67}$$

Analogously, $C^k(I;X)$, $k \in \mathbb{N}$, is defined as the space of all functions $u : I \to X$ whose derivatives in I, i.e., with respect to t, up to order k are of class $C(I;X)$.

Moreover, the Lebesgue spaces $L^p(\Omega)$, $1 \leq p < \infty$, can be generalized to $L^p(I;X)$ by substituting the absolute value in their definition and the definition of their norms by the norm on X, thus, yielding the norm

$$\|u\|_{L^p(I;X)} = \left(\int_{t \in I} \|u(t)\|_X^p \ dt \right)^{\frac{1}{p}} \ . \tag{2.68}$$

Again, $L^2(I;X)$ is a Hilbert space. The space $L^\infty(I;X)$ consists of all measurable, essentially bounded functions $u : I \to X$. It is a Banach space with respect to the norm

$$\|u\|_{L^\infty(I;X)} = \operatorname*{ess\,sup}_{t \in I} \|u(t)\|_X \ . \tag{2.69}$$

The spaces $C(\bar{I};X)$ and $L^p(\bar{I};X)$, $1 \leq p \leq \infty$, are defined accordingly.

The generalization of Sobolev spaces to X-valued functions is straight-forward. We show how this is done for $H^1(I;L^2(\Omega))$, where $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, is an open bounded domain. The corresponding norm is given by

$$\|u\|_{H^1(I;L^2(\Omega))} = \left(\int_I \left(\int_\Omega \left(|u(x,t)|^2 + |\partial_t u(x,t)|^2 \right) \mathrm{d}V(x) \right) dt \right)^{\frac{1}{2}} \ . \tag{2.70}$$

For Hilbert spaces X, the following imbedding theorem can be established:

Lemma 2.43 (Sobolev Lemma for Hilbert Space Valued Functions) *Let $I \subset \mathbb{R}$ be a bounded open interval and X be a Hilbert space. Then any function $u \in H^1(I;X)$ corresponds to a function in $C(\bar{I};X)$.*

Proof Analogous to [292, Theorem 6.2].

Remark 2.44 See, e.g., [253, Chapter III, Proposition 1.2] for a proof in a more general setting with Banach spaces instead of Hilbert spaces.

There is one more result which we use throughout this thesis regarding weak derivatives and difference quotients.

Lemma 2.45 (Weak Derivatives and Difference Quotients) *Let $\Omega \subset \mathbb{R}^n$ be open, $(0, t_{\text{end}}) \subset \mathbb{R}$, $1 < p \leq \infty$ and*

(I) $w \in \mathcal{L}^p(\Omega)$, $k \in \{1, \dots, n\}$ or
(II) $w \in L^p((0, t_{\text{end}}), X)$ with X being a Sobolev space on the set Ω.

Then the following assertions are true:

(a) If the weak partial derivatives satisfy $\partial_{x_k} w \in \mathcal{L}^p(\Omega)$ for some k in case (I) or $\partial_t w \in L^p((0, t_{\text{end}}), X)$ in case (II), we have for the corresponding difference quotients

(i) $\delta_{k,h} w \in \mathcal{L}^p(\Omega')$ for any compact subset Ω' of Ω, $h < \operatorname{dist}(\Omega', \partial\Omega)$, and $\|\delta_{k,h} w\|_{\mathcal{L}^p(\Omega')} \leq \|\partial_{x_k} w\|_{\mathcal{L}^p(\Omega)}$,
(ii) $\delta_{t,h} w \in L^p((a,b), X)$ for any compact subset $[a,b]$ of $(0, t_{\text{end}})$, $h < \min(a, t_{\text{end}} - b)$, and $\|\delta_{t,h} w\|_{L^p((a,b),X)} \leq \|\partial_t w\|_{L^p((0,t_{\text{end}}),X)}$.

This also holds for $p = 1$.

(b) (i) If there exists a constant $C \in \mathbb{R}^+$ such that for any compact subset $\Omega' \subset \Omega$ and $h < \operatorname{dist}(\Omega', \partial\Omega)$ we have $\delta_{k,h} w \in \mathcal{L}^p(\Omega')$ and $\|\delta_{k,h} w\|_{\mathcal{L}^p(\Omega')} \leq C$, then $\partial_{x_k} w$ (which is a priori well defined as a distribution) is in fact in $\mathcal{L}^p(\Omega)$ and satisfies $\|\partial_{x_k} w\|_{\mathcal{L}^p(\Omega)} \leq C$. Moreover, when $h \to 0$, $\delta_{k,h} w \to \partial_{x_k} w$ in $\mathcal{L}^p_{\text{loc}}(\Omega)$.
(ii) If there exists a constant C such that for any $[a,b] \Subset (0, t_{\text{end}})$ and $h < \min(a, t_{\text{end}} - b)$ we have $\delta_{t,h} w \in L^p((a,b), X)$ and $\|\delta_{t,h} w\|_{L^p((a,b),X)} \leq C$, then $\partial_t w$ (which is a priori well defined as a distribution) is in fact in $L^p((a,b), X)$ and satisfies $\|\partial_t w\|_{L^p((a,b),X)} \leq C$. Moreover, when $h \to 0$, $\delta_{t,h} w \to \partial_t w$ in $L^p_{\text{loc}}((a,b), X)$.

Proof See [234, Lemmata 8.48 and 8.49], as well as [113, Lemmata 8.1 and 8.2] for case (I). Case (II) can be proven analogously.

2.4 Integral Calculus

In this section, we summarize some theorems from vector calculus about integrals on vector fields. In order to do this, we have to explain some more differential operators.

We already introduced the gradient in Remark 2.9 and the corresponding notation that uses the so-called nabla operator ∇. The definition of the gradient can be generalized to differentiable vector-valued functions $u = (u_1, u_2, u_3)^T : \mathbb{R}^3 \to \mathbb{R}^3$ by defining

$$(\nabla_x u)_{ij}(x) = \partial_{x_j} u_i(x), \quad i, j \in \{1, 2, 3\} , \tag{2.71}$$

for cartesian coordinates x_i. Thus, the gradient of a vector-valued function u is identical to the transpose of the Jacobian of u.

Other often used differential operators are the divergence of u, given by

$$\nabla_x \cdot u(x) = \sum_{i=1}^{3} \partial_{x_i} u_i(x) , \tag{2.72}$$

and the curl of u, given by

$$\nabla_x \wedge u(x) = \begin{pmatrix} \partial_{x_2} u_3(x) - \partial_{x_3} u_2(x) \\ \partial_{x_3} u_1(x) - \partial_{x_1} u_3(x) \\ \partial_{x_1} u_2(x) - \partial_{x_2} u_1(x) \end{pmatrix} . \tag{2.73}$$

If a function $u : \mathbb{R}^3 \to \mathbb{R}$ is differentiable in x and its gradient $\nabla_x u$ is also differentiable in x, we can calculate the divergence of the gradient. This is known as the Laplacian, given by

$$\nabla_x^2 u(x) = \nabla_x \cdot ((\nabla_x u)(x)) = \sum_{i=1}^{3} \partial_{x_i}^2 u_i(x) . \tag{2.74}$$

The Laplacian can be defined analogously for sufficiently differentiable vector-valued functions. Note that for a vector-valued function u the expression $\nabla_x \cdot ((\nabla_x u)(x))$ differs from $\nabla_x ((\nabla_x \cdot u)(x))$.

Remark 2.46

(i) We will omit the index x in ∇_x whenever it is clear with respect to which variable the derivative is taken.
(ii) It is not necessary for the application of any of the above differential operators that the function under consideration is defined on the whole space $\mathbb{R}^3$. It may instead be defined only on an open subset $\Omega \subset \mathbb{R}^3$.
(iii) Except for the curl operator, all of the above operators can be defined not only for functions on $\mathbb{R}^3$ but $\mathbb{R}^n$, $n \in \mathbb{N}$, in a similar fashion.

With the above notation, we can formulate the following integral relations.

Theorem 2.47 (Gauß (Divergence) Theorem) *Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, $n \geq 2$ be a bounded domain with Lipschitz boundary $\partial\Omega$. If the vector field $u : \overline{\Omega} \to \mathbb{R}^n$ is an element of $\left(C^1(\overline{\Omega})\right)^n$, we have*

$$\int_\Omega \nabla_x \cdot u(x)\, \mathrm{d}V(x) = \int_{\partial\Omega} u(x) \cdot n(x)\, \mathrm{d}S(x) \tag{2.75}$$

where $\mathrm{d}V$ $(= \mathrm{d}V_n)$ denotes the volume element in $\mathbb{R}^n$, $n(x)$ denotes the outer (unit) normal field on $\partial\Omega$, and $\mathrm{d}S$ $(= \mathrm{d}S_{n-1})$ denotes the surface element of $\mathbb{R}^n$.

Proof See, e.g., [68, §3.3] in a more general context.

As a consequence of the Gauß Theorem, we get by choosing $u = v\nabla w$ with $v \in C^1(\overline{\Omega})$, $w \in C^2(\overline{\Omega})$ the following theorem.

Theorem 2.48 (First Green Theorem) *Let $\Omega \subset \mathbb{R}^3$ be a bounded domain on which the Gauß Theorem 2.47 is valid for all $u \in \mathcal{C}^1(\Omega)$. For $v \in C^1(\overline{\Omega})$ and $w \in C^2(\overline{\Omega})$, the relation*

$$\int_\Omega \left(v(x)\nabla_x^2 w(x) + (\nabla_x v(x)) \cdot (\nabla_x w(x))\right)\, \mathrm{d}V(x) = \int_{\partial\Omega} v(x)\partial_{n(x)} w(x)\, \mathrm{d}S(x) \tag{2.76}$$

holds.

Proof See, e.g., [97, Theorem 2.1].

If we insert $u = v\nabla w - w\nabla v$ with $v, w \in C^2(\overline{\Omega})$ into (2.75), we obtain

Theorem 2.49 (Second Green Theorem) *Let $\Omega \subset \mathbb{R}^3$ be a bounded domain on which the Gauß Theorem 2.47 is valid for all $u \in \mathcal{C}^1(\Omega)$. For $v, w \in C^2(\overline{\Omega})$, the relation*

$$\begin{aligned} &\int_\Omega \left(v(x)\nabla_x^2 w(x) - w(x)\nabla_x^2 v(x)\right)\, \mathrm{d}V(x) \\ &= \int_{\partial\Omega} \left(v(x)\partial_{n(x)} w(x) - w(x)\partial_{n(x)} v(x)\right)\, \mathrm{d}S(x) \end{aligned} \tag{2.77}$$

holds.

Proof See, e.g., [97, Theorem 2.2].

As we are interested in time-dependent problems, there is one more integral identity which we need. For this, we first have to define a motion [188, Chapter 1, Definitions 1.2, 1.4, and 1.5].

Definition 2.50 (Motion, Configuration, Velocity) Let $\Omega \subset \mathbb{R}^3$ be an open domain. A configuration of Ω is a mapping $\phi : \Omega \to \mathbb{R}^3$. A reference configuration

is a certain fixed configuration of Ω. Points in Ω are called material points (more often material coordinates) and denoted by $X = (X_1, X_2, X_3)^T$. Points in $\mathbb{R}^3$ are called spatial points (more often spatial coordinates) and denoted by $x = (x_1, x_2, x_3)^T$.

A motion of Ω is a mapping $\phi : \Omega \times (a, b) \to \mathbb{R}^3, (X, t) \mapsto \phi(X, t) = x$ with an open interval $(a, b) \subset \mathbb{R}$. For $t \in (a, b)$ fixed, we write $\phi_t(X)$. For a subset $\mathcal{B} \subset \Omega$, we write $\mathcal{B}_t = \phi_t(\mathcal{B})$.

A motion ϕ_t is called regular or invertible if each $\phi_t(\Omega)$ is open and the inverse $\phi_t^{-1} : \Omega_t \to \Omega$ exists. ϕ_t is called $\mathcal{C}^k$-regular, $k \in \mathbb{N}_0$, if $\phi \in \mathcal{C}^k(\Omega \times (a, b))$ and $\phi^{-1} \in \mathcal{C}^k(\Omega_t \times (a, b))$.

The material velocity $\mathcal{V}_t(X) = \mathcal{V}(X, t)$ is defined by

$$\mathcal{V}_t(X) = \partial_t \phi(X, t) \tag{2.78}$$

if the derivative exists. For $\mathcal{C}^1$-regular motions, the spatial velocity $v_t(x)$ is defined by

$$v_t : \Omega_t \to \mathbb{R}^3, \quad v_t(x) = \mathcal{V}_t\big(\phi_t^{-1}(x)\big) \ . \tag{2.79}$$

We can now define how to calculate the absolute time derivative $\frac{\mathrm{d}}{\mathrm{d}t}$ of the spatial integral of some function u.

Theorem 2.51 (Reynolds Transport Theorem) *Let $\Omega \subset \mathbb{R}^3$, $(a, b) \subset \mathbb{R}$ be an open interval and ϕ_t be a $\mathcal{C}^1$-regular motion. Let $u \in C^1(\Omega_t \times (a, b))$ and $\mathcal{B} \subset \Omega$ be open with a piecewise C^1-regular boundary $\partial\mathcal{B}$. Then*

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{B}_t} u(x, t) \, \mathrm{d}V(x) &= \int_{\mathcal{B}_t} \partial_t u(x, t) + \nabla_x \cdot \left(u(x, t) v(x, t)\right) \, \mathrm{d}V(x) \\ &= \int_{\mathcal{B}_t} \frac{\mathrm{D}u}{\mathrm{D}t}(x, t) + u(x, t) \left(\nabla_x \cdot v(x, t)\right) \, \mathrm{d}V(x) \ , \end{aligned} \tag{2.80}$$

where $\frac{\mathrm{d}}{\mathrm{d}t}$ denotes the total derivative with respect to time, ∂_t denotes the partial derivative with respect to time and

$$\frac{\mathrm{D}u}{\mathrm{D}t}(x, t) = \partial_t u(x, t) + \left(\nabla_x u(x, t)\right) \cdot v(x, t) \tag{2.81}$$

is the so-called material time derivative.

Proof See, e.g., [188, Chapter 2, Theorem 1.1] and [11, page 3].

Remark 2.52

(i) Using the correct velocity v, the one as defined in Definition 2.50, and a reference configuration $\mathcal{B}$ is crucial here. If, for example, in the settings of fluid mechanics, the fluid is allowed to leave the volume $\mathcal{B}_t$, the velocity that has to be used in the Transport Theorem may differ from the velocity of the fluid.

(ii) Applying the Gauß Theorem 2.47 to (2.80) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{B}_t} u(x,t)\,\mathrm{d}V(x) = \int_{\mathcal{B}_t} \partial_t u(x,t)\,\mathrm{d}V(x) + \int_{\partial\mathcal{B}_t} (u(x,t)v(x,t))\cdot n(x)\,\mathrm{d}S(x)\,. \tag{2.82}$$

2.5 Differential Equations

Assume we have an open bounded domain $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, $n > 1$, and a map

$$\mathscr{F}:\ \mathbb{R}^{n^k} \times \mathbb{R}^{n^{k-1}} \times \ldots \times \mathbb{R}^n \times \mathbb{R} \times \Omega \to \mathbb{R}, \quad k \in \mathbb{N}\,. \tag{2.83}$$

Then

$$\mathscr{F}\left(\left(D^k u\right)(x), \ldots, u(x), x\right) = 0 \text{ for all } x \in \Omega \tag{2.84}$$

is a partial differential equation (PDE) of order k if at least one derivative of order k is actually a part of the equation and no derivative of higher order than k is present. This can be done analogously for systems of differential equations. We only deal with linear PDEs here that can be written as

$$\sum_{|\gamma|\le k} a_\gamma(x)\left(D^\gamma u\right)(x) = f(x),\ \gamma \in \mathbb{N}_0^n\,, \tag{2.85}$$

with given coefficient functions $a_\gamma(x)$ and right-hand side $f(x)$. If $f \equiv 0$, the PDE is called homogeneous.

Within this thesis, we will deal with systems of linear PDEs of second order. There are three main classes of these PDEs. We start by defining an elliptic differential operator (cf., e.g., [72, Section 6.1.1.]).

Definition 2.53 (Elliptic PDEs) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, $u : \overline{\Omega} \to \mathbb{R}$. Let L be a linear differential operator of second order such that

$$u(x) \mapsto Lu(x) = \sum_{i,j=1}^{n} a_{ij}(x)\partial_{x_i}\partial_{x_j}u(x) + \sum_{i=1}^{n} b_i(x)\partial_{x_i}u(x) + c(x)u(x)\,. \tag{2.86}$$

L is called uniformly elliptic if there exists a constant $C > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \ge C\,\|\xi\|^2 \tag{2.87}$$

for almost every $x \in \Omega$ and all $\xi \in \mathbb{R}^n$.

With this definition, we can now define the two other main classes (cf., e.g., [72, Sections 7.1.1. and 7.1.2.] or [101, Section 1.1]). For both, there is one distinguished variable, denoted by t rather than as a component of a vector x, which is usually the time being distinct from spatial variables summarized in x.

Definition 2.54 (Parabolic PDE) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, $(0, t_{\text{end}}) \subset \mathbb{R}$, $t_{\text{end}} > 0$, $u : \overline{\Omega} \times [0, t_{\text{end}}] \to \mathbb{R}$. Let L be a linear differential operator of second order such that

$$u(x,t) \mapsto Lu(x,t) = \sum_{i,j=1}^{n} a_{ij}(x,t)\partial_{x_i}\partial_{x_j}u(x,t) + \sum_{i=1}^{n} b_i(x,t)\partial_{x_i}u(x,t) + c(x,t)u(x,t) - \partial_t u(x,t) \ . \tag{2.88}$$

L is called uniformly parabolic if there exist positive constants $\overline{\lambda}_0$, $\overline{\lambda}_1$ such that

$$\overline{\lambda}_0 \|\xi\|^2 \leq \sum_{i,j=1}^{n} a_{ij}(x,t)\xi_i\xi_j \leq \overline{\lambda}_1 \|\xi\|^2 \tag{2.89}$$

for all $(x,t) \in \Omega \times (0, t_{\text{end}})$ and all $\xi \in \mathbb{R}^n$.

Definition 2.55 (Hyperbolic PDE) Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, $(0, t_{\text{end}}) \subset \mathbb{R}$, $t_{\text{end}} > 0$, $u : \overline{\Omega} \times [0, t_{\text{end}}] \to \mathbb{R}$. Let L be a linear differential operator of second order such that

$$u(x,t) \mapsto Lu(x,t) = \sum_{i,j=1}^{n} a_{ij}(x,t)\partial_{x_i}\partial_{x_j}u(x,t) + \sum_{i=1}^{n} b_i(x,t)\partial_{x_i}u(x,t) + c(x,t)u(x,t) - \partial_t^2 u(x,t) \ . \tag{2.90}$$

L is called uniformly hyperbolic if there exists a constant $C > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x,t)\xi_i\xi_j \geq C \|\xi\|^2 \tag{2.91}$$

for all $(x,t) \in \Omega \times (0, t_{\text{end}})$ and all $\xi \in \mathbb{R}^n$.

Remark 2.56

(i) It is possible that the character of a differential operator changes with x, e.g., when there is a function of x as coefficient of the time derivative term. Such equations can be locally elliptic, parabolic, or hyperbolic instead of uniformly, i.e., they are of one of these types on certain subdomains of Ω.

(ii) Not all linear second order PDEs are of one of the above classes for $n > 2$. Indeed, the system of equations which we discuss in this thesis is of neither of the above types, but has similarities with some elliptic and parabolic equations.

We are interested in finding solutions of differential equations. Usually, we have some more information than just the differential equation itself in the form of initial and boundary conditions. From a theoretical point of view, these often help to guarantee uniqueness of solutions, or at least specify in which way solutions to the same problem may differ. Instead of discussing this in general, we will treat this problem in the next chapters for the set of equations in which we are interested in this thesis.

As already pointed out, the formulations of differential equations as given above with strong partial derivatives is often not suited to find answers to the questions of solvability, uniqueness of solutions, or their regularity. Instead, we would like to have a formulation based on weak derivatives.

Let us assume that a linear second order differential equation is given in its strong form by

$$Lu(x) = f(x) \text{ for all } x \in \Omega \ . \tag{2.92}$$

For simplicity, we equip this equation with the homogeneous Dirichlet boundary condition

$$u(x) = 0 \text{ for all } x \in \partial\Omega \ . \tag{2.93}$$

In the context of differential equations, we will often use the abbreviation $\Gamma = \partial\Omega$.

A classical strong solution of this PDE has to be in $C^2(\Omega)$ which is a rather restrictive requirement. We can relax this requirement in two points. First, we can get to a weak differentiable solution. For this purpose, suppose v is an arbitrary function belonging to $C_0^\infty(\Omega)$, multiply the differential equation by v and integrate over Ω. We obtain

$$(Lu, v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \text{ for all } v \in C_0^\infty(\Omega) \ . \tag{2.94}$$

Moreover, we can relax the requirements on differentiability of u by performing an integration by parts on the left-hand side, which gives us a bilinear form $a(u, v)$. Thus, a useful assumption on u is $u \in H_0^1(\Omega)$. Additionally, as $C_0^\infty(\Omega)$ is a dense subspace of $H_0^1(\Omega)$, we can extend the space of functions with which we multiply to $H_0^1(\Omega)$. This yields

$$a(u, v) = f(v) \text{ for all } v \in H_0^1(\Omega) \ . \tag{2.95}$$

Here, we interpreted the right-hand side as a linear functional on $H_0^1(\Omega)$. It is easy to see that every solution of the strong formulation is also a solution of the weak formulation. However, the opposite may not be true.

For other kinds of boundary conditions, the above procedure is changed in two points. On the one hand, if the values of u on the boundary are given and different from zero, u has to belong to another subspace of $H^1(\Omega)$. On the other hand, if normal derivatives of u are specified in a Neumann boundary condition, integration by parts yields some integrals over (parts of) the boundary Γ of Ω. These are usually incorporated into the linear form f on the right-hand side. Other modifications may be necessary for other boundary conditions.

The answer to the question, whether a solution of (2.95) exists, depends on the properties of $a(\cdot\,,\cdot)$.

Definition 2.57 (Coercitivity) Let V be a real Hilbert space with inner product $(\cdot\,,\cdot)_V$ and corresponding norm $\|\cdot\|_V$.

A bilinear form $a(\cdot\,,\cdot) : V \times V \to \mathbb{R}$ is strict coercive if a constant $C > 0$ exists such that

$$a(u,u) \geq C\,\|u\|_V^2 \tag{2.96}$$

for all $u \in V$.

If $a(\cdot\,,\cdot)$ is a continuous coercive bilinear form and f a continuous linear functional on V, the task to determine $u \in V$ such that we have

$$a(u,v) = f(v) \text{ for every } v \in V \tag{2.97}$$

is called a variational problem. It is equivalent to determining $u \in V$ such that

$$\frac{1}{2}a(u,u) - f(u) = \min_{0 \neq v \in V} \left(\frac{1}{2}a(v,v) - f(v) \right) . \tag{2.98}$$

For a proof of this equivalence, see, e.g., [55, Theorem 1.1.2].

Under the additional assumptions that $a(\cdot\,,\cdot)$ is symmetric and positive definite, i.e., it is an inner product, solvability of the variational problem is given by the Riesz Representation Theorem.

Theorem 2.58 (Riesz Representation Theorem) *Let V be a Hilbert space and $a(\cdot\,,\cdot)$ be an inner product on V. For every continuous linear functional $f \in V'$, there is a unique $u \in V$ such that (2.97) is valid.*

Proof See, e.g., [295, Section III.6].

For non-symmetric bilinear forms, we have

Theorem 2.59 (Lax-Milgram Theorem) *Let V be a Hilbert space and $a(\cdot\,,\cdot)$ be a continuous and strict coercive bilinear form on $V \times V$. For every continuous linear functional $f \in V'$, there is a unique $u \in V$ such that (2.97) is valid.*

Proof See, e.g., [295, Section III.7].

In problems in which a special role is assigned to time, a different setting, in which time is regarded as a parameter, is advantageous [175, Chapter IV]. Moreover, we consider two Hilbert spaces V and $\mathscr{H}$ such that V is densely imbedded in $\mathscr{H}$, $\mathscr{H}' = \mathscr{H}$, and, thus, $V \subseteq \mathscr{H} \subseteq V'$. The pair $(V, \mathscr{H})$ is called a rigged Hilbert space or Gelfand triple. For $t \in (0, t_{\text{end}})$, $t_{\text{end}} > 0$, we consider a family $a(t; \cdot, \cdot)$ of continuous bilinear forms on V and assume that for $u, v \in V$ the function $t \mapsto a(t; u, v)$ is measurable and that there is a constant $C \in \mathbb{R}^+$, independent of t, u, and v, such that (cf. [175, Chapter IV, (1.1)])

$$|a(t; u, v)| \leq C \|u\|_V \|v\|_V \ . \tag{2.99}$$

Using such a family of bilinear forms, the weak formulation of a parabolic PDE may be given as the following problem [175, Chapter IV, Problem 1.1]: Find $u \in L^2((-\infty, t_{\text{end}}), V)$ such that $u(t) = 0$ for almost every $t < 0$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} (u(t), v)_V + a(t; u(t), v) = (f(t), v)_V + (u^0, v)_V \delta(t) \tag{2.100}$$

with $f \in L^2((-\infty, t_{\text{end}}), \mathscr{H})$ vanishing for $t < 0$ and u^0 is given in $\mathscr{H}$. For this problem, we have the following theorem.

Theorem 2.60 (Lions) *Let V, $\mathscr{H}$ be as defined at the beginning of this paragraph and let $a(t; \cdot, \cdot)$ be a family of continuous bilinear forms on V for $t \in [0, t_{\text{end}}]$ with (2.99). If $a(t; \cdot, \cdot)$ satisfies*

$$a(t; v, v) \geq c \|v\|_V - \beta \|v\|_{\mathscr{H}} \text{ for all } v \in V \tag{2.101}$$

with some constant $c \in \mathbb{R}^+$ and $\beta \in \mathbb{R}$, there exists a unique solution u of the above given problem and the map $\{f, u^0\} : L^2((-\infty, t_{\text{end}}), \mathscr{H}) \times \mathscr{H} \to V$ is continuous.

The condition (2.101) is called modified coercitivity or Gårding inequality.

Proof See, e.g., [175, Chapter IV, Theorem 1.1].

Chapter 3
Physical and Mathematical Foundation

This chapter is supposed to fulfill two purposes. We start by giving a short introduction into poroelasticity from a physical point of view and deduce the partial differential equations of quasistatic poroelasticity starting from physical balance equations and using constitutive equations first given by Biot [34, 35]. In a second section, we address the question of existence and uniqueness of solutions.

A shortened version of this chapter was already published before as part of a technical report [13], two articles in peer-reviewed journals [14, 17], and a book chapter [15].

3.1 Physical Background

This section explains the physical basics of linear elasticity in general and especially linear poroelasticity. Our presentation is mainly based on the books by Landau et al. [170], Lai et al. [169], and Jaeger et al. [138], as well as the work by Auriault [18], Philips and Wheeler [221, 222], and an article by Showalter [254].

3.1.1 Linear Elasticity

Consider a domain $\Omega \subset \mathbb{R}^3$. Points in Ω can either be considered with respect to Ω as material points or material coordinates, denoted by X, or with respect to $\mathbb{R}^3$ as spatial points or spatial coordinates, denoted by x. Both descriptions are coupled, as the spatial coordinates can be seen as a motion of the material coordinates, i.e., $x = \phi(X, t)$ for some $t \in (0, t_{\text{end}})$, $t_{\text{end}} > 0$. If the inverse of the motion exists, we

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4_3

can also write $X = \phi^{-1}(x,t)$. This leads to the definition of the displacement vector

$$u(X,t) = x(X,t) - X \ . \tag{3.1}$$

Let us consider a point at an infinitesimal distance to x, i.e, $x + \mathrm{d}x$. For such a point, we get

$$x + \mathrm{d}x = X + \mathrm{d}X + u(X + \mathrm{d}X, t) \ . \tag{3.2}$$

Using (3.1) and a first-order Taylor series for $u(X + \mathrm{d}X, t)$, we obtain

$$\begin{aligned} \mathrm{d}x =& \mathrm{d}X + u(X + \mathrm{d}X, t) + X - x \\ \approx& \mathrm{d}X + u(X,t) + \frac{\partial u(X,t)}{\partial X_k}\mathrm{d}X_k - u(X,t) \\ =& \mathrm{d}X + \frac{\partial u(X,t)}{\partial X_k}\mathrm{d}X_k \ . \end{aligned} \tag{3.3}$$

Here and afterwards, we use Einstein's summation convention. From a physical point of view, it is also of interest how the length of $\mathrm{d}x$ can be expressed by $\mathrm{d}X$ and u. We obtain

$$\begin{aligned} \|\mathrm{d}x\|^2 - \|\mathrm{d}X\|^2 =& \mathrm{d}x_i \mathrm{d}x_i - \mathrm{d}X_j \mathrm{d}X_j \\ =& \frac{\partial x_i}{\partial X_k}\mathrm{d}X_k \frac{\partial x_i}{\partial X_l}\mathrm{d}X_l - \frac{\partial X_j}{\partial X_k}\mathrm{d}X_k \frac{\partial X_j}{\partial X_l}\mathrm{d}X_l \\ =& \left(\frac{\partial x_i}{\partial X_k}\frac{\partial x_i}{\partial X_l} - \delta_{jk}\delta_{jl}\right)\mathrm{d}X_k \mathrm{d}X_l \\ =& \left(\frac{\partial x_i}{\partial X_k}\frac{\partial x_i}{\partial X_l} - \delta_{kl}\right)\mathrm{d}X_k \mathrm{d}X_l \ . \end{aligned} \tag{3.4}$$

With (3.1), we get

$$\begin{aligned} \|\mathrm{d}x\|^2 - \|\mathrm{d}X\|^2 &= \left(\frac{\partial\left(X_i + u_i(X,t)\right)}{\partial X_k}\frac{\partial\left(X_i + u_i(X,t)\right)}{\partial X_l} - \delta_{kl}\right)\mathrm{d}X_k \mathrm{d}X_l \\ &= \left(\left(\delta_{ik} + \frac{\partial u_i}{\partial X_k}\right)\left(\delta_{il} + \frac{\partial u_i}{\partial X_l}\right) - \delta_{kl}\right)\mathrm{d}X_k \mathrm{d}X_l \\ &= \left(\frac{\partial u_l}{\partial X_k} + \frac{\partial u_k}{\partial X_l} + \frac{\partial u_i}{\partial X_k}\frac{\partial u_i}{\partial X_l}\right)\mathrm{d}X_k \mathrm{d}X_l \ . \end{aligned} \tag{3.5}$$

This relation leads to the definition of the Lagrangian strain tensor

$$\boldsymbol{\mathcal{E}}_{ij}(u(X,t)) = \frac{1}{2}\left(\frac{\partial u_i}{\partial X_j}(X,t) + \frac{\partial u_j}{\partial X_i}(X,t) + \frac{\partial u_k}{\partial X_i}(X,t)\frac{\partial u_k}{\partial X_j}(X,t)\right) \tag{3.6}$$

which is a symmetric tensor of rank 2. Similarly, the Eulerian (or Euler-Almansi) strain tensor can be defined as

$$\boldsymbol{e}_{ij}(u(x,t)) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}(x,t) + \frac{\partial u_j}{\partial x_i}(x,t) - \frac{\partial u_k}{\partial x_i}(x,t)\frac{\partial u_k}{\partial x_j}(x,t)\right) . \tag{3.7}$$

The difference between these two strains is in whether we use material coordinates or spatial coordinates, which naturally yields the difference in sign of the last summand in each tensor.

Within this thesis, we follow the usual assumption that only small deformations are considered, such that u and derivatives of u are small. This implies that the difference between material coordinates and spatial coordinates is small, such that either description should yield the same expressions. Due to the smallness of u and its derivatives, products of derivatives of u are negligible, leading to the definition of the infinitesimal or Cauchy strain tensor

$$\boldsymbol{\epsilon}_{ij}(u(x,t)) = \frac{1}{2}\left(\partial_{x_j} u_i(x,t) + \partial_{x_i} u_j(x,t)\right) , \tag{3.8}$$

which replaces either the Lagrangian or Eulerian strain tensor. The corresponding framework to describe deformations is called linear elasticity or infinitesimal strain theory.

Remark 3.1 A strict derivation of the theory of linear elasticity as linearization of a more general non-linear theory needs a preparation involving differential geometry, differential forms and notation of co- and contravariant vectors, tensors and derivatives. Thus, it is beyond the scope of this thesis. The interested reader is referred to [188]. There is also an argumentation based on considerations from physics in [56] that explains why elasticity can never be linear, but only approximated by a linear theory. Nevertheless, the linearized theory is the one most commonly used.

The average of the strain tensor's diagonal elements is called the volumetric strain ϵ_v and can be computed by

$$3\epsilon_v = \operatorname{tr}(\boldsymbol{\epsilon}). \tag{3.9}$$

It is linked to the change of volume due to compression of a body.

By deforming a body, inner forces, so-called inner stresses, arise. These inner stresses are surface forces in contrast to body forces like gravitation. The inner stresses can be characterized by the second rank (Cauchy) stress tensor $\boldsymbol{\sigma}$. To derive balance equations for $\boldsymbol{\sigma}$, we consider an arbitrary material volume $\mathcal{B}_t \subset \Omega$ with boundary $\partial\mathcal{B}_t$. Let v_s be the velocity with which $\mathcal{B}_t$ deforms. Then, conservation of linear momentum yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{B}_t} \rho(x,t) v_s(x,t)\, \mathrm{d}V(x) = \int_{\mathcal{B}_t} f(x,t)\, \mathrm{d}V(x) + \int_{\partial\mathcal{B}_t} \boldsymbol{\sigma}(x,t) n(x)\, \mathrm{d}S(x). \tag{3.10}$$

Here, f denotes a body force density, ρ is the mass density, $n(x)$ is the outer normal unit vector to $\partial\mathcal{B}_t$ at a point x of the boundary, and $\frac{\mathrm{d}}{\mathrm{d}t}$ denotes the total derivative with respect to time.

Applying Reynolds Transport Theorem 2.51 to the left-hand side and Gauß Theorem 2.47 to the right-hand side of (3.10), we get

$$\begin{aligned}&\int_{\mathcal{B}_t}\left(\frac{\mathrm{D}(\rho(x,t)v_s(x,t))}{\mathrm{D}t}+\rho(x,t)v_s(x,t)\,(\nabla_x\cdot v_s(x,t))\right)\mathrm{d}V(x)\\&=\int_{\mathcal{B}_t}\left(f(x,t)+\nabla_x\cdot\boldsymbol{\sigma}(x,t)\right)\mathrm{d}V(x)\end{aligned}\tag{3.11}$$

with the material time derivative $\frac{\mathrm{D}}{\mathrm{D}t}$ as defined in Theorem 2.51. Moreover,

$$\begin{aligned}\frac{\mathrm{D}(\rho(x,t)v_s(x,t))}{\mathrm{D}t}&=\frac{\mathrm{D}\rho(x,t)}{\mathrm{D}t}v_s(x,t)+\rho(x,t)\frac{\mathrm{D}v_s(x,t)}{\mathrm{D}t}\\&=-\left(\rho(x,t)\nabla_x\cdot v_s(x,t)\right)v_s(x,t)+\rho(x,t)\frac{\mathrm{D}v_s(x,t)}{\mathrm{D}t}\,,\end{aligned}\tag{3.12}$$

where we applied conservation of mass (cf. [11] and [188, Chapter 1, Theorem 5.7]), given by

$$\partial_t\rho(x,t)+\nabla_x\cdot\left(\rho(x,t)v_s(x,t)\right)=\frac{\mathrm{D}\rho(x,t)}{\mathrm{D}t}+\rho(x,t)\left(\nabla_x\cdot v_s(x,t)\right)=0\,.\tag{3.13}$$

This yields

$$\int_{\mathcal{B}_t}\rho(x,t)\frac{\mathrm{D}v_s(x,t)}{\mathrm{D}t}\,\mathrm{d}V(x)=\int_{\mathcal{B}_t}\left(f(x,t)+\nabla_x\cdot\boldsymbol{\sigma}(x,t)\right)\mathrm{d}V(x)\,,\tag{3.14}$$

from which we get by using the arbitrariness of $\mathcal{B}_t$ and assuming continuity of the integrand the balance of linear momentum as

$$\rho(x,t)\frac{\mathrm{D}v_s(x,t)}{\mathrm{D}t}=f(x,t)+\nabla_x\cdot\boldsymbol{\sigma}(x,t)\,.\tag{3.15}$$

The velocity $v_s(x,t)$ is still unknown, but it can be connected to the displacement $u(x,t)$ by

$$v_s(x,t)=\frac{\mathrm{D}x(t)}{\mathrm{D}t}=\frac{\mathrm{D}u(x,t)}{\mathrm{D}t}=\partial_t u(x,t)+v_s(x,t)\left(\nabla_x\cdot u(x,t)\right)\,,\tag{3.16}$$

which is an implicit equation for $v_s(x,t)$. As mentioned before, we assume that the displacement vector $u(x,t)$ changes slowly. Therefore, also the velocity $v_s(x,t)$ has to be small. Hence, the second term on the right-hand side can be neglected. The same reasoning stays true for derivatives of $v_s(x,t)$ with respect to time which give us the balance of linear momentum as

$$\rho(x,t)\partial_t^2 u(x,t) = f(x,t) + \nabla_x \cdot \boldsymbol{\sigma}(x,t) \tag{3.17}$$

since $\mathcal{B}_t$ is arbitrary.

Similarly, balance of angular momentum is given by

$$\begin{aligned}
&\int_{\mathcal{B}_t} x(t) \wedge \left(\rho(x,t)\ \partial_t^2 u(x,t)\right)\ \mathrm{d}V(x) \\
&= \int_{\mathcal{B}_t} x(t) \wedge f(x,t)\ \mathrm{d}V(x) + \int_{\partial\mathcal{B}_t} x(t) \wedge (\boldsymbol{\sigma}(x,t)n(x))\ \mathrm{d}S(x)\ .
\end{aligned} \tag{3.18}$$

Rewriting the right-hand side of (3.18) by using the Levi-Civita tensor $\boldsymbol{\varepsilon}_{ijk}$ and applying the Gauß Theorem 2.47 to the second integral yields

$$\begin{aligned}
&\int_{\mathcal{B}_t} \boldsymbol{\varepsilon}_{ijk} x_j(t) f_k(x,t)\ \mathrm{d}V(x) + \int_{\partial\mathcal{B}_t} \boldsymbol{\varepsilon}_{ijk} x_j(t) \boldsymbol{\sigma}_{kl}(x,t) n_l(x,t)\ \mathrm{d}S(x) \\
&= \int_{\mathcal{B}_t} \boldsymbol{\varepsilon}_{ijk} x_j(t) f_k(x,t)\ \mathrm{d}V(x) + \int_{\mathcal{B}_t} \partial_{x_l} \left(\boldsymbol{\varepsilon}_{ijk} x_j(t) \boldsymbol{\sigma}_{kl}(x,t)\right)\ \mathrm{d}V(x) \\
&= \int_{\mathcal{B}_t} \boldsymbol{\varepsilon}_{ijk} x_j(t) f_k(x,t)\ \mathrm{d}V(x) \\
&\quad + \int_{\mathcal{B}_t} \boldsymbol{\varepsilon}_{ijk} \Bigg(\underbrace{\underbrace{\partial_{x_l} x_j(t)}_{=\delta_{jl}} \boldsymbol{\sigma}_{kl}(x,t)}_{=\boldsymbol{\sigma}_{kj}(x,t)} + x_j(t) \underbrace{\partial_{x_l} \boldsymbol{\sigma}_{kl}(x,t)}_{=\rho(x,t)\partial_t^2 u_k(x,t) - f_k(x,t)} \Bigg)\ \mathrm{d}V(x)\ .
\end{aligned} \tag{3.19}$$

Assuming continuity of $\boldsymbol{\sigma}$ and taking into account that $\mathcal{B}_t$ is an arbitrary volume, we get

$$\boldsymbol{\varepsilon}_{ijk} \boldsymbol{\sigma}_{kj} = 0\ . \tag{3.20}$$

Thus, $\boldsymbol{\sigma}$ is a symmetric tensor (cf. [11, Equations (5.3), (5.4)]). As for the strain tensor $\boldsymbol{\epsilon}$, the average of the diagonal elements of $\boldsymbol{\sigma}$ is of special interest. It is called the mean normal stress σ_m and can be computed by

$$3\sigma_m = \operatorname{tr}(\boldsymbol{\sigma}) \ . \tag{3.21}$$

One of the objects of elasticity theory is to find a relation between stress $\boldsymbol{\sigma}$ and strain $\boldsymbol{\epsilon}$. As mentioned before, we only consider small deformations. Therefore, it is sufficient to assume a linear relation between stress and strain which can be written as

$$\boldsymbol{\sigma}_{ij}(x,t) = C_{ijkl}(x,t)\boldsymbol{\epsilon}_{kl}(u(x,t)) \ . \tag{3.22}$$

Here, C_{ijkl} is the so-called (Cauchy) elasticity tensor of rank 4, which is symmetric because of the symmetry of $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$.

Throughout this thesis, we will only consider isotropic homogeneous materials and time-independent elastic constants. Consequently, the elasticity tensor can be written as

$$C_{ijkl}(x,t) = \lambda\delta_{ij}\delta_{kl} + \mu\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) \text{ for all } (x,t) \ . \tag{3.23}$$

Here, λ and μ are the so-called Lamé coefficients. Usually, it is required that $\mu > 0$ and $3\lambda + 2\mu > 0$. For isotropic materials, (3.22) becomes

$$\boldsymbol{\sigma}_{ij}(x,t) = \lambda\boldsymbol{\epsilon}_{kk}(u(x,t))\delta_{ij} + 2\mu\boldsymbol{\epsilon}_{ij}(u(x,t)) \ . \tag{3.24}$$

Summing up the diagonal elements on both sides gives

$$3\sigma_m = (3\lambda + 2\mu)\,\epsilon_v = 3K\epsilon_v \ . \tag{3.25}$$

Here, we introduced the bulk modulus K. It is one of the elastic moduli among which are also the shear modulus G, Young's modulus E, and the Poisson ratio ν. The elastic moduli are not independent from each other and are also linked to the Lamé coefficients. The relations between them are [221, appendix A.1]

$$G = \mu \ , \tag{3.26a}$$

$$K = \lambda + \frac{2}{3}\mu \ , \tag{3.26b}$$

$$E = \mu\frac{2\mu + 3\lambda}{\mu + \lambda} \ , \tag{3.26c}$$

$$\nu = \frac{\lambda}{2\,(\mu + \lambda)} \ . \tag{3.26d}$$

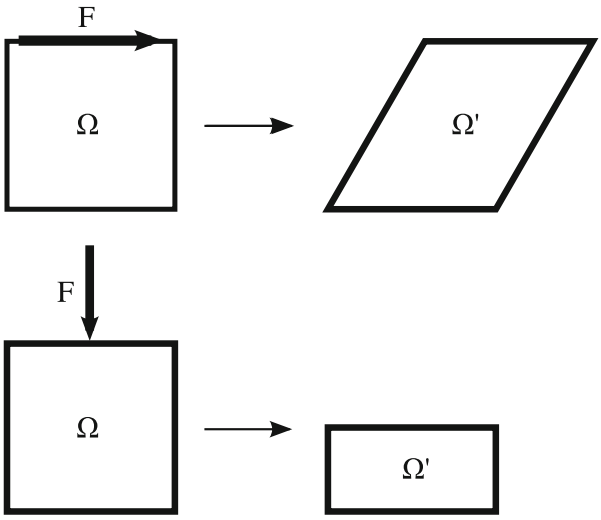

Fig. 3.1 Visualization of stresses and how they work on an elastic body Ω. *Top*: Shearing a domain Ω by applying a tangential force F changes the orientation of surfaces but not the volume. The lower part of the boundary of Ω is assumed to be fixed. *Bottom*: Compressing the domain Ω by applying a normal force F changes the volume of the body but not the orientation of surfaces. The lower and lateral parts of boundary of Ω are assumed to be fixed

Thus, by measuring the values of two of these six variables, the other four can be determined. Another useful quantity is the compressibility C which is the inverse of the bulk modulus K.

A 2D-visualization of how stresses may work on a body is given in Fig. 3.1. The top row shows how a tangentially applied force F shears the domain Ω. The bottom row shows how a normal force F compresses the domain Ω. Note that pure compression changes the volume of a body and preserves the orientation of surfaces, while pure shearing preserves the volume but changes the orientation of surfaces.

3.1.2 Poroelasticity

The term poroelasticity was first used by J. Geertsma in 1966 [285] to describe the modeling of coupled mechanics and flow in porous media. The first model to describe the influence of fluid flow on the deformation of soils was developed by von Terzaghi in 1923 [271], who invented a one-dimensional model for consolidation. A consistent theory for three-dimensional linear poroelasticity was developed by Biot [34–36, 38]. Additionally to the stress $\boldsymbol{\sigma}$ discussed in the last section, we now have to consider the fluid pore pressure p. According to Biot, a constitutive equation for the total poroelastic stress $\boldsymbol{\sigma}^{\mathrm{pe}}$ is given by

$$\boldsymbol{\sigma}^{\mathrm{pe}}(x,t) = \boldsymbol{\sigma}(x,t) - \alpha \boldsymbol{I} p(x,t) \ . \tag{3.27}$$

Here, $\boldsymbol{I}$ is the unity tensor of rank 2 and α is the so-called Biot-Willis constant which can be expressed as [65]

$$\alpha = 1 - \frac{K}{K_m}, \quad 0 < \alpha \leq 1 \,, \tag{3.28}$$

where K is the bulk modulus of the total medium consisting of rock and fluid and K_m is the bulk modulus of the rock matrix.

In most cases, the pore pressure has to be considered unknown. Thus, we need another equation to determine it. In order to find such an equation, we borrow the argumentation of Jaeger et al. [138] regarding the changes in volume. Here, we introduce three volumes – the bulk volume V_b, the rock matrix volume V_m, and the pore volume V_p – which satisfy

$$V_b = V_m + V_p \tag{3.29}$$

and give rise to the definition of porosity ϕ by

$$\phi = \frac{V_p}{V_b} \,. \tag{3.30}$$

To characterize the reaction of the medium to compression, four compressibilities can be defined:

$$C_{b\sigma} = \frac{1}{V_b^{(i)}} \left(\frac{\partial V_b}{\partial \sigma_m^{\mathrm{pe}}} \right)_p \,, \quad C_{bp} = \frac{1}{V_b^{(i)}} \left(\frac{\partial V_b}{\partial p} \right)_{\sigma_m^{\mathrm{pe}}} \,, \tag{3.31a}$$

$$C_{p\sigma} = \frac{1}{V_p^{(i)}} \left(\frac{\partial V_p}{\partial \sigma_m^{\mathrm{pe}}} \right)_p \,, \quad C_{pp} = \frac{1}{V_p^{(i)}} \left(\frac{\partial V_p}{\partial p} \right)_{\sigma_m^{\mathrm{pe}}} \,. \tag{3.31b}$$

The superscript "(i)" denotes the initial, unstressed state and the notation $(\ldots)_X$ for derivatives implies that the derivative is taken while the quantity X is fixed. σ_m^{pe} is given by $\sigma_m^{\mathrm{pe}} = \sigma_m - \alpha p$. The compressibilities satisfy

$$C_m = C_{p\sigma} - C_{pp} \,, \tag{3.32a}$$

$$C_m = C_{b\sigma} - C_{bp} \,, \tag{3.32b}$$

$$C_{bp} = \phi C_{p\sigma} \,, \tag{3.32c}$$

$$C_{b\sigma} = \phi C_{pp} + (1 + \phi) C_m \,, \tag{3.32d}$$

where $C_m = \frac{1}{K_m}$ is the compressibility of the rock matrix.

To connect the change of volume with the pressure, we note that V_p can be written as

$$V_p = \frac{m_f}{\rho_f} \tag{3.33}$$

with the mass of the fluid, m_f, and its mass density, ρ_f, assuming that the medium is saturated, i.e., the pores are completely filled up with fluid. Changes of the pore volume are then given by

$$\mathrm{d}V_p = \mathrm{d}\left(\frac{m_f}{\rho_f}\right) = \frac{\mathrm{d}m_f}{\rho_f} - \frac{m_f \mathrm{d}\rho_f}{\rho_f^2} = \frac{\mathrm{d}m_f}{\rho_f} - \underbrace{\frac{m_f}{\rho_f}}_{=V_p} \frac{\mathrm{d}\rho_f}{\rho_f} = \frac{\mathrm{d}m_f}{\rho_f} - V_p C_f \mathrm{d}p \,. \tag{3.34}$$

Here, we defined the fluid compressibility C_f via $\frac{\mathrm{d}\rho_f}{\rho_f} = C_f \mathrm{d}p$. Dividing all terms by the bulk volume gives

$$\frac{\mathrm{d}V_p}{V_b} = \frac{1}{V_b}\frac{\mathrm{d}m}{\rho_f} - \phi C_f \mathrm{d}p \,. \tag{3.35}$$

Thus, the change in the fluid content can be divided into one part due to compression or expansion of the fluid (second term on the right-hand side) and another part solely due to mass transfer (first term on the right-hand side) which we define as $\mathrm{d}\zeta$. Using the compressibilities (3.32), we get

$$\mathrm{d}\zeta = \frac{1}{V_b}\frac{\mathrm{d}m}{\rho_f} = \frac{\mathrm{d}V_p}{V_b} + \phi C_f \mathrm{d}p = \phi\left[C_{p\sigma}\mathrm{d}\sigma_m^{\mathrm{pe}} + \left(C_{pp} + C_f\right)\mathrm{d}p\right] \,. \tag{3.36}$$

Recalling Definition (3.28) of α, we get

$$\phi C_{p\sigma} = \frac{\alpha}{K} \,. \tag{3.37}$$

Moreover, we define Skempton's coefficient B by

$$B = \frac{C_{p\sigma}}{C_{pp} + C_f} \tag{3.38}$$

and use $\epsilon_v = \frac{\sigma_m}{K}$ to get

$$\mathrm{d}\zeta = \frac{\alpha}{K}\mathrm{d}\sigma_m^{\mathrm{pe}} + \frac{\alpha}{KB}\mathrm{d}p = \alpha \mathrm{d}\epsilon_v + \frac{\alpha}{K}\left(\frac{1}{B} - \alpha\right)\mathrm{d}p \,. \tag{3.39}$$

This yields the definition of the constrained specific storage coefficient c_0:

$$c_0 = \frac{\alpha}{K}\left(\frac{1}{B} - \alpha\right) \,. \tag{3.40}$$

Because of the independence of the constitutive parameters of stress (and, thus, strain) and pressure in linearized poroelasticity, we can integrate (3.39). Note that per definition $\zeta = 0$ in the unstressed state and we get

$$\zeta(x,t) = c_0 p(x,t) + \alpha \nabla_x \cdot u(x,t) \ . \tag{3.41}$$

Here, we used $\epsilon_v(x,t) = \nabla_x \cdot u(x,t)$.

Now the equation we are looking for can be derived from conservation of mass which implies

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{B}_t} \zeta(x,t) \, \mathrm{d}V(x) = \int_{\mathcal{B}_t} h(x,t) \, \mathrm{d}V(x) \ . \tag{3.42}$$

Although this form of the conservation of mass seems to involve only volumes, it is valid because ζ is defined as that portion of volume change which is solely due to mass transfer while compressibility effects are accounted separately. Again, applying the Reynolds Transport Theorem 2.51 and assuming continuity of all quantities as well as arbitrariness of $\mathcal{B}_t$ yields

$$\partial_t \zeta(x,t) + \nabla_x \cdot v_f(x,t) = h(x,t) \ , \tag{3.43}$$

where v_f is the fluid velocity.

A relation between the fluid velocity v_f and the pore pressure p was first found empirically by Darcy [63] and is given in a more general form by [31]

$$v_f(x,t) = -\frac{1}{\mu_f} \boldsymbol{\kappa}(x,t) \left(\nabla_x p(x,t) - g(x,t) \right) \ . \tag{3.44}$$

Here, $\boldsymbol{\kappa}$ is the permeability tensor, μ_f is the fluid viscosity, and g is the fluid body force.

Combining Eqs. (3.8), (3.17), (3.24), (3.27), (3.41), (3.43), and (3.44) yields

$$\rho(x,t)\partial_t^2 u(x,t) - (\lambda + \mu) \nabla_x \left(\nabla_x \cdot u(x,t) \right) - \mu \nabla_x^2 u(x,t) + \alpha \nabla_x p(x,t) = f(x,t) \ , \tag{3.45a}$$

$$\partial_t \left(c_0 p(x,t) + \alpha \nabla_x \cdot u(x,t) \right) - \nabla_x \cdot \left(\boldsymbol{k}(x,t) \left(\nabla_x p(x,t) - g(x,t) \right) \right) = h(x,t) \tag{3.45b}$$

as the governing equations of poroelasticity. Here, we assumed μ_f to be constant and set $\boldsymbol{k} = \frac{1}{\mu_f} \boldsymbol{\kappa}$. Note that in (3.17), instead of $\boldsymbol{\sigma}$ the total stress $\boldsymbol{\sigma}^{\mathrm{pe}}$ has to be used to deduce these equations.

In most applications, the term $\rho(x,t)\partial_t^2 u(x,t)$ in (3.45a) is neglected. To motivate this, we follow an argument by Phillips [221, Appendix B] and non-dimensionalize (3.45). For simplicity, we assume homogeneous permeability such that $\boldsymbol{k}$ can be

replaced by a scalar constant k and a constant mass density ρ. Let x_0, t_0 be a characteristic lengthscale and a characteristic timescale, respectively. Setting

$$\check{x} = \frac{x}{x_0}\,, \quad \check{t} = \frac{t}{t_0}\,,$$
$$\check{u} = \frac{u}{x_0}\,, \quad \check{p} = \frac{p}{\mu}\,, \quad \check{f} = \frac{x_0}{\mu} f\,, \quad \check{g} = \frac{x_0}{\mu} g\,, \quad \check{h} = t_0 h\,,$$

we obtain

$$\left[\frac{\rho x_0^2}{\mu t_0^2}\right] \partial_{\check{t}}^2 \check{u} - \left[\frac{\lambda+\mu}{\mu}\right] \check{\nabla}_{\check{x}} \left(\check{\nabla}_{\check{x}} \cdot \check{u}\right) - \check{\nabla}_{\check{x}}^2 \check{u} + \alpha \check{\nabla}_{\check{x}} \check{p} = \check{f}\,, \tag{3.46a}$$

$$\partial_{\check{t}} \left(c_0 \mu \check{p} + \alpha \left(\check{\nabla}_{\check{x}} \cdot \check{u}\right)\right) - \left[\frac{t_0 k \mu}{x_0^2}\right] \check{\nabla}_{\check{x}}^2 \check{p} + \left[\frac{t_0 k \mu}{x_0^2}\right] \check{\nabla}_{\check{x}} \cdot \check{g} = \check{h}\,. \tag{3.46b}$$

For the sake of readability, we omitted the dependencies on x, t, $\check{x}$ and $\check{t}$. The dimensionless coefficients in front of $\check{\nabla}_{\check{x}}^2 \check{p}$ and $\check{\nabla}_{\check{x}} \check{g}$ in (3.46b) suggest

$$t_0 = \frac{x_0^2}{k\mu}\,. \tag{3.47}$$

Plugging in this value into (3.46) yields

$$\left[\frac{\rho \mu k^2}{x_0^2}\right] \partial_{\check{t}}^2 \check{u} - \left[\frac{\lambda+\mu}{\mu}\right] \check{\nabla}_{\check{x}} \left(\check{\nabla}_{\check{x}} \cdot \check{u}\right) - \check{\nabla}_{\check{x}}^2 \check{u} + \alpha \check{\nabla}_{\check{x}} \check{p} = \check{f}\,, \tag{3.48a}$$

$$\partial_{\check{t}} \left(c_0 \mu \check{p} + \alpha \left(\check{\nabla}_{\check{x}} \cdot \check{u}\right)\right) - \check{\nabla}_{\check{x}}^2 \check{p} + \check{\nabla}_{\check{x}} \cdot \check{g} = \check{h}\,. \tag{3.48b}$$

As long as we are interested in consolidation processes, x_0 can be assumed to be of the order of several hundred meters or several kilometers. Thus, the factor in front of $\partial_{\check{t}}^2 \check{u}$ becomes rather small. Assuming $x_0 = 100$ m and Berea sandstone [247] as material, we get

$$\frac{\rho \mu k^2}{x_0^2} \approx 5.3 \cdot 10^{-11}\,,$$

where

$$\frac{\lambda+\mu}{\mu} = \frac{5}{3}\,, \ \alpha = 0.867\,, \text{ and } c_0 \mu \approx 0.461\,.$$

Therefore, the term $\rho(x,t)\partial_t^2 u(x,t)$ in (3.45a) is negligible, whereas the time derivative of the pressure in (3.48b) must not be neglected. However, using the

characteristic length $x_0 = k\sqrt{\rho\left(\lambda + 2\mu + \frac{\alpha^2}{c_0}\right)}$ as suggested for the study of wave propagation (see [247] and the references therein), we obtain

$$\frac{\rho\mu k^2}{x_0^2} = \frac{\mu}{\lambda + 2\mu + \frac{\alpha^2}{c_0}} \approx 0.2 \, .$$

This much larger value confirms that the second derivative of u with respect to time must not be neglected when studying poroelastic waves.

Throughout this thesis, we will only be interested in consolidation processes and, thus, use the **quasistatic equations of poroelasticity (QEP)**

$$-(\lambda + \mu)\nabla_x(\nabla_x \cdot u) - \mu\nabla_x^2 u + \alpha\nabla_x p = f \, , \tag{3.49a}$$

$$\partial_t(c_0 p + \alpha\nabla_x \cdot u) - \nabla_x \cdot (\boldsymbol{k}(\nabla_x p - g)) = h \, . \tag{3.49b}$$

Here and further on, we omit the dependency on x and t and will mention it explicitly only if it is necessary for understanding or to avoid confusion.

Besides the above introduced coefficients, other moduli used in poroelasticity are Biot's modulus M, the undrained bulk modulus K_u, the undrained Poisson's ratio ν_u, and the fluid diffusivity tensor $\boldsymbol{c}_f$. These are related to the ones we already introduced by

$$M = \frac{1}{c_0} \, , \tag{3.50a}$$

$$K_u = K + \frac{\alpha^2}{c_0} \, , \tag{3.50b}$$

$$\nu_u = \frac{\lambda + \frac{\alpha^2}{c_0}}{2\left(\mu + \lambda + \frac{\alpha^2}{c_0}\right)} \, , \tag{3.50c}$$

$$\boldsymbol{c}_f = \frac{\lambda + 2\mu}{c_0(\lambda + 2\mu) + \alpha^2}\boldsymbol{k} \, . \tag{3.50d}$$

In order to find a unique solution of (3.49), boundary and initial conditions are needed. One possibility to choose boundary conditions is

$$u(x,t) = u_D(x,t) \quad \text{for all } (x,t) \in \Gamma_d \times [0, t_{\text{end}}] \, , \tag{3.51a}$$

$$\boldsymbol{\sigma}^{\text{pe}}(x,t)n(x) = t_n(x,t) \quad \text{for all } (x,t) \in \Gamma_t \times [0, t_{\text{end}}] \, , \tag{3.51b}$$

$$p(x,t) = p_D(x,t) \quad \text{for all } (x,t) \in \Gamma_p \times [0, t_{\text{end}}] \, , \tag{3.51c}$$

$$-\boldsymbol{k}(\nabla_x p(x,t) - g(x,t)) \cdot n(x) = v_{f,n}(x,t) \quad \text{for all } (x,t) \in \Gamma_f \times [0, t_{\text{end}}] \, , \tag{3.51d}$$

where $\overline{\Gamma_d} \cup \overline{\Gamma_t} = \partial\Omega$ and $\overline{\Gamma_p} \cup \overline{\Gamma_f} = \partial\Omega$ are two different disjoint partitions of the domain boundary $\partial\Omega = \Gamma$, $n(x)$ is the outward normal to $\partial\Omega$, t_n is the normal tension, and $v_{f,n}$ is the normal fluid velocity. The indices "d", "t", "p", and "f" indicate whether displacement, tension, pressure, or fluid velocity is prescribed on the boundary. In the next section, we will discuss which initial conditions are suitable for (3.49).

Remark 3.2 The equations above, which we deduced for poroelasticity, are the same as the equations for pure thermoelasticity, if the pressure is replaced by the temperature T and the coefficients are adapted (see [254] for details).

Remark 3.3 The dynamic equations of poroelasticity are also used to model the behavior of cerebrospinal fluid pressure and parenchymal displacement in the brain [291].

3.2 Existence and Uniqueness in Quasistatic Poroelasticity

The aim of this section is to present results on the existence and uniqueness of solutions of the system of differential Eqs. (3.49) together with the boundary conditions in (3.51) and suitable initial conditions when transferred into a weak formulation instead of the strong formulation of the equations presented in the previous section. Such results are given by Auriault and Sanchez-Palencia [19] as well as by Showalter et al. who used different boundary conditions and also discussed several related systems in a series of papers [254–259]. The results presented here are mainly taken from [19, 254] and adapted to the notation used within this thesis with some additional notes.

Remark 3.4 We assume $\alpha \neq 0$, $c_0 \neq 0$, and $k \neq 0$ are scalar and constant. If $\alpha = 0$, (3.49) decouples into the Cauchy-Navier equation of classical linear elasticity theory (see below) and an equation similar to the heat equation [200]. For those two differential equations, existence of unique solutions is known. The case $c_0 = 0$ is discussed by Ženíšek [301]. If $k = 0$, the second equation could be solved by integration with respect to time, then substitute p into the first equation yielding some modified equation of elasticity. The existence and uniqueness of solutions for this case is shown in [217].

First of all, we derive a weak formulation of (3.49) together with the boundary conditions in (3.51). For the sake of simplicity, we neglect the fluid body force g in (3.49b). It can be considered as part of the right-hand side term h. Moreover, we assume homogeneous Dirichlet boundary conditions. Inhomogeneous Dirichlet boundary values can be treated similarly.

Let us define two function spaces

$$V = \left\{ v \in \mathcal{H}^1(\Omega) : v|_{\Gamma_d} = 0 \right\} , \tag{3.52a}$$

$$Q = \left\{ q \in H^1(\Omega) : q|_{\Gamma_p} = 0 \right\} \tag{3.52b}$$

with norms $\|\cdot\|_V$ and $\|\cdot\|_Q$ induced by the norm on $\mathcal{H}^1(\Omega)$ and $H^1(\Omega)$, respectively. We assume $u \in V$, $p \in Q$.

Please note that throughout this section, v is used to denote test functions which are always in V and should neither be confused with the deformation velocity v_s of an elastic solid, nor with the fluid velocity v_f which we used in the previous section.

Using the total poroelastic stress from (3.27), Eq. (3.49a) can be written as

$$-\partial_{x_j} \boldsymbol{\sigma}_{ij}^{\mathrm{pe}} = f_i \, . \tag{3.53}$$

Multiplying this equation with a test function $v \in V$ and integrating by parts over the domain Ω yields

$$\begin{aligned} -\int_{\Omega} \left(\partial_{x_j} \boldsymbol{\sigma}_{ij}^{\mathrm{pe}} \right) v_i \, \mathrm{d}V(x) &= \int_{\Omega} \boldsymbol{\sigma}_{ij}^{\mathrm{pe}} \left(\partial_{x_j} v_i \right) \, \mathrm{d}V(x) - \int_{\partial\Omega} \boldsymbol{\sigma}_{ij}^{\mathrm{pe}} v_i n_j(x) \, \mathrm{d}S(x) \\ &= \int_{\Omega} f \cdot v \, \mathrm{d}V(x) \, . \end{aligned} \tag{3.54}$$

The boundary integral can be split into an integral over Γ_d which vanishes because $v \in V$ and an integral over Γ_t for which we use the Neumann boundary condition from (3.51b). Due to the symmetry of $\boldsymbol{\sigma}_{ij}^{\mathrm{pe}}$, we can use the strain $\boldsymbol{\epsilon}_{ij}(v)$ corresponding to v instead of $\partial_{x_j} v_i$. Using (3.27) and (3.22) we get

$$\begin{aligned} \int_{\Omega} C_{ijkl} \boldsymbol{\epsilon}_{kl}(u) \boldsymbol{\epsilon}_{ij}(v) \, \mathrm{d}V(x) &= \int_{\Omega} \alpha \delta_{ij} p \boldsymbol{\epsilon}_{ij}(v) \, \mathrm{d}V(x) \\ &\quad + \int_{\Omega} f \cdot v \, \mathrm{d}V(x) + \int_{\Gamma_t} t_n \cdot v \, \mathrm{d}S(x) \, . \end{aligned} \tag{3.55}$$

Note, that in the last integral on the right-hand side, the normal $n(x)$ is contained in t_n. The integral on the left-hand side is well known from the theory of linear elasticity. It defines a symmetric bilinear form $e(\cdot, \cdot)$ which is continuous

with respect to $\|\cdot\|_{\mathcal{H}^1(\Omega)}$. Continuity results from the Cauchy-Schwarz inequality. Moreover, we have the inequality

$$\begin{aligned} e(v,v) &= \int_\Omega C_{ijkl}\boldsymbol{\epsilon}_{kl}(v)\boldsymbol{\epsilon}_{ij}(v)\,\mathrm{d}V(x) \\ &= \int_\Omega \left(\lambda\boldsymbol{\epsilon}_{jj}(v)\boldsymbol{\epsilon}_{kk}(v) + 2\mu\boldsymbol{\epsilon}_{ij}(v)\boldsymbol{\epsilon}_{ij}(v)\right)\,\mathrm{d}V(x) \\ &\geq 2\mu \int_\Omega \boldsymbol{\epsilon}_{ij}(v)\boldsymbol{\epsilon}_{ij}(v)\,\mathrm{d}V(x) \end{aligned} \tag{3.56}$$

for isotropic materials with $\lambda > 0$, $\mu > 0$. To show that $e(\cdot\,,\cdot)$ is coercive, we need

Theorem 3.5 (Korn's Inequality) *Let Ω be a domain in $\mathbb{R}^3$. For each $v \in \mathcal{H}^1(\Omega)$, let*

$$\boldsymbol{\epsilon}(v) = \frac{1}{2}\left(\partial_{x_j} v_i + \partial_{x_i} v_j\right) \in \boldsymbol{L}^2(\Omega)\ .$$

Then there exists a constant $c > 0$ such that

$$\|v\|_{\mathcal{H}^1(\Omega)} \leq c\left(\|v\|^2_{\mathcal{L}^2(\Omega)} + \|\boldsymbol{\epsilon}(v)\|^2_{\boldsymbol{L}^2(\Omega)}\right)^{\frac{1}{2}}\ , \tag{3.57}$$

and, thus, the mapping

$$v \mapsto \left(\|v\|^2_{\mathcal{L}^2(\Omega)} + \|\boldsymbol{\epsilon}(v)\|^2_{\boldsymbol{L}^2(\Omega)}\right)^{\frac{1}{2}} \tag{3.58}$$

is a norm on $\mathcal{H}^1(\Omega)$, equivalent to $\|\cdot\|_{\mathcal{H}^1(\Omega)}$.

Proof See, e.g., [102].

From Korn's inequality, the coercivity of $e(\cdot\,,\cdot)$ results by

Theorem 3.6 *Let Ω be a domain in $\mathbb{R}^3$, let Γ_d be a measurable subset of $\partial\Omega$. Then the space V defined in (3.52a) is a closed subspace of $\mathcal{H}^1(\Omega)$. If $\mathrm{meas}(\Gamma_d) > 0$, there exists a constant $c > 0$ such that*

$$c^{-1}\,\|v\|_{\mathcal{H}^1(\Omega)} \leq \|\boldsymbol{\epsilon}(v)\|_{\boldsymbol{L}^2(\Omega)} \leq c\,\|v\|_{\mathcal{H}^1(\Omega)} \text{ for all } v \in V\ , \tag{3.59}$$

i.e., on the space V, the semi-norm $v \mapsto \|\boldsymbol{\epsilon}(v)\|_{\boldsymbol{L}^2(\Omega)}$ is a norm, equivalent to $\|\cdot\|_{\mathcal{H}^1(\Omega)}$.

Proof See [56, Theorem 6.3-4].

Putting together (3.56) and Theorem 3.6, we get

$$|e(v,v)| \geq 2\mu \left| \int_\Omega \boldsymbol{\epsilon}_{ij}(v)\boldsymbol{\epsilon}_{ij}(v)\, \mathrm{d}V(x) \right| = 2\mu \, \|\boldsymbol{\epsilon}(v)\|^2_{\boldsymbol{L}^2(\Omega)} \geq C \, \|v\|_{\mathcal{H}^1(\Omega)} \tag{3.60}$$

with some constant $C > 0$.

Remark 3.7

(i) The assumption $\lambda > 0$, $\mu > 0$ can be weakened to $\lambda > -\frac{2}{3}\mu$, $\mu > 0$.
(ii) For conditions on coercivity of anisotropic materials, see [127].

In order to solve (3.55), we first solve the auxiliary problem

$$e(u_{\mathrm{ce}}, v) = \int_\Omega f \cdot v \, \mathrm{d}V(x) + \int_{\Gamma_t} t_n \cdot v \, \mathrm{d}S(x) \tag{3.61}$$

for each $t \geq 0$ in order to eliminate the inhomogeneous term on the right-hand side and inhomogeneous Neumann boundary conditions. This is a weak formulation of the so-called Cauchy-Navier equation, the classical problem of linear elasticity. Thus, we have

Theorem 3.8 (Existence and Uniqueness in Classical Linear Elasticity) *Let Ω be a domain in $\mathbb{R}^3$, let Γ_d be a measurable subset of $\partial\Omega$ with* $\mathrm{meas}(\Gamma_d) > 0$*, let $f \in \mathcal{L}^2(\Omega)$, $t_n \in \mathcal{H}^{-\frac{1}{2}}(\Gamma_t)$. Then there is one and only one function $u_{\mathrm{ce}} \in V$ that satisfies the Cauchy-Navier Equation (3.61).*

If, moreover, $\overline{\Gamma_d} \cap \overline{\Gamma_t} = \emptyset$, $t_n \in \mathcal{H}^{\frac{1}{2}}(\Gamma_t)$, and $\partial\Omega$ has $\mathcal{C}^2$-regularity, u_{ce} is in $\mathcal{H}^2(\Omega)$.

Proof Continuity of $\int_\Omega f \cdot v \, \mathrm{d}V(x) + \int_{\Gamma_t} t_n \cdot v \, \mathrm{d}S(x)$ on V is obvious for $f \in \mathcal{L}^2(\Omega)$ and results from the Trace Theorem 2.41 for $t_n \in \mathcal{H}^{-\frac{1}{2}}(\Gamma_t)$. Coercivity of $e(\cdot,\cdot)$ was proven in Theorem 3.6. Thus, existence and uniqueness of u_{ce} result from the Lax-Milgram Theorem 2.59.

For regularity of u_{ce}, see [56, Theorem 6.3-6] and the references therein.

Remark 3.9

(i) In case of inhomogeneous Dirichlet boundary values, we still solve the auxiliary problem with the homogeneous condition $u_{\mathrm{ce}} = 0$ on Γ_d.
(ii) For further results on linear elasticity (e.g., slightly more general conditions on f and t_n, pure Dirichlet boundary conditions, pure Neumann boundary conditions, higher regularity), see also [56, Chapter 6] as well as [193] and the references therein.

Now we replace u by $u = \hat{u} + u_{\mathrm{ce}}$ and get from (3.55) the equivalent set of Eqs. (3.61) and

$$\int_{\Omega} C_{ijkl}\boldsymbol{\epsilon}_{kl}(\hat{u})\boldsymbol{\epsilon}_{ij}(v)\mathrm{d}V(x) = \int_{\Omega} \alpha\, \delta_{ij}\, p\, \boldsymbol{\epsilon}_{ij}(v)\, \mathrm{d}V(x)\ . \tag{3.62}$$

Note that this implies a homogeneous Neumann condition in (3.51b) on Γ_t. The term on the right-hand side of (3.62) is a continuous linear functional on $\mathcal{H}^1(\Omega)$ because

$$\int_{\Omega} \alpha\, \delta_{ij}\, p\, \boldsymbol{\epsilon}_{ij}(v)\, \mathrm{d}V(x) \leq c\, \|p\|_{L^2(\Omega)}\, \|v\|_{\mathcal{H}^1(\Omega)}\ . \tag{3.63}$$

Thus, Theorem 3.8 and (3.62) enable the definition of a unique $\hat{u}$ as a function of p not only for $p \in Q$, but also on the larger space $p \in L^2(\Omega)$. Therefore, the operator

$$p \mapsto \hat{u}(p)$$

is linear and continuous from $L^2(\Omega)$ to V. The question of existence and uniqueness of solutions of (3.49) can be reduced to existence and uniqueness of a solution p of (3.49b). However, note that due to the substitution $u = \hat{u} + u_{\mathrm{ce}}$, the right-hand side of (3.49b) has to be modified to

$$\hat{h} = h - \alpha \partial_t\, (\nabla_x \cdot u_{\mathrm{ce}})\ . \tag{3.64}$$

Thus, any condition made on $\hat{h}$ concerning to which function space $\hat{h}$ belongs has to be satisfied by $\partial_t\, (\nabla_x \cdot u_{\mathrm{ce}})$, too, which leads to a stronger condition on u_{ce}. We will discuss this later when we know to which space $\hat{h}$ has to belong.

To handle the coupling term in (3.49b), we consider that $\hat{u}$ can be defined as a function of p. Therefore, it is possible to replace $\hat{u}$ in (3.49b) by defining the operator B via

$$Bp = \alpha \nabla_x \cdot \hat{u}\ . \tag{3.65}$$

This operator is linear and continuous from $L^2(\Omega)$ to $L^2(\Omega)$. By using $v = \hat{u}(q)$ as a test function in (3.62) and observing that $\nabla_x \cdot \hat{u} = \delta_{ij}\boldsymbol{\epsilon}_{ij}(\hat{u})$ we get

$$\begin{aligned}(Bq, p)_{L^2(\Omega)} &= \int_{\Omega} \alpha \delta_{ij} \boldsymbol{\epsilon}_{ij}(\hat{u}(q))\, p\, \mathrm{d}V(x) \\ &= \int_{\Omega} C_{ijkl}\boldsymbol{\epsilon}_{kl}(\hat{u}(p))\, \boldsymbol{\epsilon}_{ij}(\hat{u}(q))\ \mathrm{d}V(x)\ .\end{aligned} \tag{3.66}$$

Thus, B is self-adjoint since C_{ijkl} is symmetric. As a consequence we have

$$(Bp, p)_{L^2(\Omega)} \geq 0 \,. \tag{3.67}$$

Applying the First Green Theorem 2.48, the second term on the left-hand side of (3.49b) leads to

$$a(p, q) = \int_{\Omega} k \nabla_x p \cdot \nabla_x q \, \mathrm{d}V(x) \,, \quad p, q \in Q \,, \tag{3.68}$$

which is a symmetric and continuous bilinear form on Q. Moreover, $a(\cdot\,, \cdot)$ satisfies the modified coercivity or Gåarding inequality

$$a(p, p) = \int_{\Omega} k \nabla_x p \cdot \nabla_x p \, \mathrm{d}V(x) \geq c \, \|p\|_{H^1(\Omega)} - \beta \, \|p\|_{L^2(\Omega)} \tag{3.69}$$

with some constants $\beta, \ c > 0$ if $k > 0$ [234, Theorem 8.26]. As before, we used integration by parts which now also yields a boundary integral term because of the non-vanishing Neumann boundary condition in (3.51d). We put this term on the right-hand side which then reads

$$\overline{h}(q) = (\hat{h} \oplus v_{f,n}, q) = \int_{\Omega} \hat{h} q \, \mathrm{d}V(x) + \int_{\Gamma_f} v_{f,n} \, q \, \mathrm{d}S(x) \,. \tag{3.70}$$

In the last integral on the right-hand side, the normal $n(x)$ is contained in $v_{f,n}$. $\overline{h}$ is a linear form on Q as long as $\hat{h} \in L^2(\Omega)$ and $v_{f,n} \in H^{-\frac{1}{2}}(\Gamma_f)$ for $t \in (0, t_{\text{end}})$ [234, Example 10.7]. Altogether, we gain the following operator formulation of (3.49b)

$$\frac{\mathrm{d}}{\mathrm{d}t}(c_0 p + Bp) + Ap = \overline{h} \tag{3.71}$$

which is an implicit evolution equation. Here, we defined the operator $A(p) : \ Q \to Q'$ via

$$(Ap, q)_{L^2(\Omega)} = a(p, q) \,. \tag{3.72}$$

For the first term on the left-hand side, we have

$$\|p\|^2_{L^2(\Omega)} \leq (c_0 p + Bp, p)_{L^2(\Omega)} \leq c \, \|p\|^2_{L^2(\Omega)} \,. \tag{3.73}$$

It defines a continuous coercive bilinear form and, because of its symmetry, even an inner product on $L^2(\Omega)$ which we refer to by

$$b(p, q) = (c_0 p + Bp, q)_{L^2(\Omega)} \,. \tag{3.74}$$

We would like to apply a result similar to Theorem 2.60. For this purpose, we have to consider some different formulations of evolution problems. According to Lions [175, Chapter I], it is useful to consider the following ones:

(a) Let $\mathfrak{H}$ be a Banach space and $\mathfrak{A}(t)$ be a family of operators in $\mathfrak{H}$ such that $\mathfrak{A}(t)$ is closed and $D(\mathfrak{A}(t)) \subset \mathfrak{H}$ dense $\forall t \geq 0$. Find $\mathfrak{u}(t)$ such that $t \mapsto \mathfrak{u}(t)$ is continuously differentiable with respect to t $\forall t \geq 0$ in $\mathfrak{H}$ with

$$\mathfrak{u}(t) \in D(\mathfrak{A}(t)) \quad \forall t \geq 0 \,, \tag{3.75a}$$

$$\mathfrak{A}(t)\mathfrak{u}(t) + \mathfrak{u}'(t) = \mathfrak{f}(t) \,, \quad \mathfrak{u}'(t) = \frac{\mathrm{d}\mathfrak{u}(t)}{\mathrm{d}t} \,, \tag{3.75b}$$

$$\mathfrak{u}(0) = \mathfrak{u}^0 \tag{3.75c}$$

for given $\mathfrak{f}(t)$ with $t \mapsto \mathfrak{f}(t)$ strongly continuous with respect to t, $\forall t \geq 0$, in $\mathfrak{H}$ and given $\mathfrak{u}^0 \in D(\mathfrak{A}(0))$.

(b) Under the same assumptions as in (a), find $\mathfrak{u} \in L^2((0, t_{\mathrm{end}}), \mathfrak{H})$ with $\mathfrak{u}' \in L^2((0, t_{\mathrm{end}}), \mathfrak{H})$ and $\mathfrak{u}(0) = \mathfrak{u}^0 \in \mathfrak{H}$ such that

$$\int_0^{t_{\mathrm{end}}} \left((\mathfrak{u}(t), \mathfrak{A}^*(t)\varphi(t))_{\mathfrak{H}} + (\mathfrak{u}'(t), \varphi(t))_{\mathfrak{H}} \right) \mathrm{d}t = \int_0^{t_{\mathrm{end}}} (\mathfrak{f}(t), \varphi(t))_{\mathfrak{H}} \, \mathrm{d}t \tag{3.76}$$

$\forall \varphi(t)$ with $\mathfrak{A}^*(t)$ being the family of adjoint operators of $\mathfrak{A}(t)$ and

(1) $t \mapsto \varphi(t)$ continuous from $[0, t_{\mathrm{end}}]$ to $\mathfrak{H}$ or from $[0, \infty)$ to $\mathfrak{H}$ if $t_{\mathrm{end}} = \infty$,
(2) $\varphi(t_{\mathrm{end}}) = 0$ or $\varphi(t) = 0$ $\forall t \in (C, \infty)$ for some constant $C \in \mathbb{R}^+$ if $t_{\mathrm{end}} = \infty$,
(3) $\varphi(t) \in D(\mathfrak{A}^*(t))$ $\forall t \in [0, t_{\mathrm{end}}]$, $t \mapsto \mathfrak{A}^*(t)\varphi(t)$ continuous from $[0, t_{\mathrm{end}}]$ to $\mathfrak{H}$.

(c) By using integration by parts, (b) yields:
Find $\mathfrak{u} \in L^2((0, t_{\mathrm{end}}), \mathfrak{H})$ such that

$$\int_0^{t_{\mathrm{end}}} \left((\mathfrak{u}(t), \mathfrak{A}^*(t)\varphi(t))_{\mathfrak{H}} - (\mathfrak{u}(t), \varphi'(t))_{\mathfrak{H}} \right) \mathrm{d}t = \int_0^{t_{\mathrm{end}}} (\mathfrak{f}(t), \varphi(t))_{\mathfrak{H}} \, \mathrm{d}t + (\mathfrak{u}^0, \varphi(0))_{\mathfrak{H}} \tag{3.77}$$

$\forall \varphi(t)$ like in (b) with $t \mapsto \varphi(t)$ continuously differentiable with respect to t from $[0, t_{\mathrm{end}}]$ to $\mathfrak{H}$.

(d) The equation can also be formulated by using bilinear forms. Consider two separable Hilbert spaces $\mathfrak{V}, \mathfrak{H}$ with $\mathfrak{V} \subset \mathfrak{H}$ topologically and algebraically and

$\mathfrak{V}$ dense in $\mathfrak{H}$. Let $\mathfrak{a}(t;\mathfrak{u},\mathfrak{v})$ be a family of bilinear forms on $\mathfrak{V}$ with parameter $t \in [0, t_{\text{end}}]$, $t_{\text{end}} < \infty$. and

(1) $t \mapsto \mathfrak{a}(t;\mathfrak{u},\mathfrak{v})$ is measurable $\forall \mathfrak{u}, \mathfrak{v} \in \mathfrak{V}$, i.e., $\mathfrak{u}$, $\mathfrak{v}$ time independent,
(2) $|\mathfrak{a}(t;\mathfrak{u},\mathfrak{v})| < M \|\mathfrak{u}\|_{\mathfrak{V}} \|\mathfrak{v}\|_{\mathfrak{V}}$ for some constant $M > 0$ independent of t, $\mathfrak{u}$, $\mathfrak{v}$.

Find $\mathfrak{u} \in L^2((-\infty, t_{\text{end}}), \mathfrak{V})$ with

$$\int_0^{t_{\text{end}}} \left(\mathfrak{a}(t;\mathfrak{u}(t),\varphi(t)) - (\mathfrak{u}(t),\varphi'(t))_{\mathfrak{H}}\right) \mathrm{d}t = \int_0^{t_{\text{end}}} (\mathfrak{f}(t),\varphi(t))_{\mathfrak{H}} \, \mathrm{d}t + (\mathfrak{u}^0,\varphi(0)) \tag{3.78}$$

$\forall \varphi \in L^2((0, t_{\text{end}}), \mathfrak{V})$ with $\varphi' \in L^2((0, t_{\text{end}}), \mathfrak{H})$, $\varphi(t_{\text{end}}) = 0$, and given $\mathfrak{f} \in L^2((0, t_{\text{end}}), \mathfrak{H})$ as well as $\mathfrak{u}^0 \in \mathfrak{H}$.

(e) Without integrals with respect to t, (d) can be written as:
Find $\mathfrak{u} \in L^2((0, t_{\text{end}}), \mathfrak{V})$ with

$$\mathfrak{u} = 0 \text{ f.a.e. } t < 0 \,, \tag{3.79a}$$

$$\mathfrak{a}(t;\mathfrak{u}(t),\mathfrak{v}) + \frac{\mathrm{d}}{\mathrm{d}t}(\mathfrak{u}(t),\mathfrak{v})_{\mathfrak{H}} = (\mathfrak{f}(t),\mathfrak{v})_{\mathfrak{H}} + (\mathfrak{u}^0,\mathfrak{v})_{\mathfrak{H}}\delta(t) \tag{3.79b}$$

$\forall \mathfrak{v} \in \mathfrak{V}$ and given $\mathfrak{f} \in L^2((-\infty, t_{\text{end}}), \mathfrak{H})$, $\mathfrak{f} = 0$ f.a.e. $t < 0$ as well as $\mathfrak{u}^0 \in \mathfrak{H}$.

Lemma 3.10 *The weak formulations (b)–(e) of (a) are equivalent.*

Proof See [175, page 2–5 and Chapter IV, Lemma 1.1 and 1.2].

Because of the operator $c_0 I + B$, Eq. (3.71) cannot directly be reformulated as in (e) but in a slightly more general form. Therefore, we first generalize formulation (a) [175, Chapter 1, Remark 1.4].

(a.1) Let $\mathfrak{B}(t)$ be another family of unbounded operators in $\mathfrak{H}$ with domain $D(\mathfrak{B})$. Find a function $t \mapsto \mathfrak{u}(t)$, continuous for $t > 0$ in $\mathfrak{H}$ with

$$\mathfrak{u}(t) \in D(\mathfrak{A}) \cap D(\mathfrak{B}) \quad \forall t \geq 0 \,. \tag{3.80a}$$

The function $t \mapsto \mathfrak{B}(t)\mathfrak{u}(t)$ is continuously differentiable for $t \geq 0$ in $\mathfrak{H}$ satisfying

$$\mathfrak{A}(t)\mathfrak{u}(t) + \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathfrak{B}(t)\mathfrak{u}(t)\right) = \mathfrak{f}(t) \text{ for } t > 0 \text{ with} \tag{3.80b}$$

$$\mathfrak{B}(0)\mathfrak{u}(0) = \mathfrak{B}(0)\mathfrak{u}^0 \text{ given in } R\left(\mathfrak{B}(0)\right) \,. \tag{3.80c}$$

In order to find a formulation of this equation similar to (e), we carry out the same steps, getting:

(e.1) Let $\mathfrak{V}_0$, $\mathfrak{V}_1$ be two Hilbert spaces with inner products $(\cdot,\cdot)_i$, $i = 0, 1$, and associated norms $\|\cdot\|_i$. Let $\mathfrak{F}$ be a locally convex topological vector space with $\mathfrak{V}_0 \subset \mathfrak{F}$ and $\mathfrak{V}_1 \subset \mathfrak{F}$ topologically and algebraically. Define the Hilbert space $\mathfrak{V} = \mathfrak{V}_0 \cap \mathfrak{V}_1$ with inner product $(\mathfrak{u},\mathfrak{v})_{\mathfrak{V}} = (\mathfrak{u},\mathfrak{v})_0 + (\mathfrak{u},\mathfrak{v})_1$. Consider two families of bilinear forms $\mathfrak{a}_i(t;\mathfrak{u},\mathfrak{v})$ continuous on $\mathfrak{V}_0$ and $\mathfrak{V}_1$, respectively, for $t \in [0, t_{\text{end}}]$ which satisfy the conditions:

(I) For $\mathfrak{u},\mathfrak{v} \in \mathfrak{V}_1$, i.e., independent of time, $t \mapsto \mathfrak{a}_1(t;\mathfrak{u},\mathfrak{v})$ is continuous on $[0, t_{\text{end}}]$ with

(1)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathfrak{a}_1(t;\mathfrak{u},\mathfrak{v})\right) = \mathfrak{a}_1'(t;\mathfrak{u},\mathfrak{v}) \text{ is measurable,} \tag{3.81a}$$

(2)

$$\left|\mathfrak{a}_1'(t;\mathfrak{u},\mathfrak{v})\right| \le M_1 \|\mathfrak{u}\|_1 \|\mathfrak{v}\|_1 \text{ with some constant } M_1 > 0\ , \tag{3.81b}$$

(3)

$$\mathfrak{a}_1(t;\mathfrak{u},\mathfrak{v}) = \overline{\mathfrak{a}_1(t;\mathfrak{v},\mathfrak{u})}\ , \tag{3.81c}$$

(4)

$$\mathfrak{a}_1(t;\mathfrak{v},\mathfrak{v}) \ge \alpha_1 \|\mathfrak{v}\|_1^2 \text{ for some constant } \alpha_1 > 0 \text{ and all } \mathfrak{v} \in \mathfrak{V}_1. \tag{3.81d}$$

(II) For $\mathfrak{u},\mathfrak{v} \in \mathfrak{V}_0$, i.e., independent of time, $t \mapsto \mathfrak{a}_0(t;\mathfrak{u},\mathfrak{v})$ is continuous on $[0, t_{\text{end}}]$ with

(1)

$$\mathfrak{a}_0(t;\mathfrak{u},\mathfrak{v}) \text{ is measurable,} \tag{3.81e}$$

(2)

$$|\mathfrak{a}_0(t;\mathfrak{u},\mathfrak{v})| \le M_0 \|\mathfrak{u}\|_0 \|\mathfrak{v}\|_0 \text{ with some constant } M_0\ . \tag{3.81f}$$

(III) There exists $\lambda \in \mathbb{R}$ such that

$$\mathrm{Re}\left(\mathfrak{a}_0(t;\mathfrak{v},\mathfrak{v})\right) + \lambda \mathfrak{a}_1(t;\mathfrak{v},\mathfrak{v}) \ge c\left(\|\mathfrak{v}\|_0^2 + \|\mathfrak{v}\|_1^2\right) \tag{3.81g}$$

for some constant $c > 0$ and all $\mathfrak{v} \in \mathfrak{V} = \mathfrak{V}_0 \cap \mathfrak{V}_1$.

We consider the following problem:
Find $\mathfrak{u} \in L^2((-\infty, t_{\text{end}}), \mathfrak{V})$ with

$$\mathfrak{u} = 0 \text{ f.a.e. } t < 0 \,, \tag{3.82a}$$

$$\mathfrak{a}_0(t; \mathfrak{u}(t), \mathfrak{v}) + \frac{\mathrm{d}}{\mathrm{d}t}(\mathfrak{a}_1(t; \mathfrak{u}(t), \mathfrak{v})) = (\mathfrak{f}(t), \mathfrak{v})_{\mathfrak{V}} + \mathfrak{a}_1(0; \mathfrak{u}^0, \mathfrak{v})\delta(t) \text{ for all } \mathfrak{v} \in \mathfrak{V} \tag{3.82b}$$

with given $\mathfrak{f} \in L^2((-\infty, t_{\text{end}}), \mathfrak{V})$, $\mathfrak{f} = 0$ f.a.e. $t < 0$, and $\mathfrak{u}^0 \in \mathfrak{V}_1$.

The existence and uniqueness of solutions of (3.82) under Conditions (3.81) is given by

Theorem 3.11 *Under Conditions* (3.81)*, there exists a unique solution* $\mathfrak{u}$ *of Problem* (3.82)*.*

The mapping $\left(\mathfrak{f}, \mathfrak{u}^0\right) \mapsto \mathfrak{u}$ *is continuous from* $L^2((-\infty, t_{\text{end}}), \mathfrak{V}) \times \mathfrak{V}_1$ *to* $L^2((-\infty, t_{\text{end}}), \mathfrak{V})$.

Proof See [175, Chapter IV, Theorem 7.1].

In case of quasistatic poroelasticity, for an appropriate weak formulation of (3.71), we have

- $\mathfrak{V}_0 \hat{=} Q$ with the inner product $(\cdot, \cdot)_0$ which is induced by the inner product $(\cdot, \cdot)_{H^1(\Omega)}$ on $H^1(\Omega)$,
- $\mathfrak{V}_1 \hat{=} L^2(\Omega)$ with the standard inner product $(\cdot, \cdot)_1 \hat{=} (\cdot, \cdot)_{L^2(\Omega)}$,
- $\Rightarrow \mathfrak{V} \hat{=} Q$ with the sum of the above mentioned inner products as inner product $(\cdot, \cdot)_{\mathfrak{V}} \hat{=} (\cdot, \cdot)_Q = (\cdot, \cdot)_{L^2(\Omega)} + (\cdot, \cdot)_{H^1(\Omega)}$,
- $\mathfrak{H} = \mathfrak{F} \hat{=} L^2(\Omega)$,
- $\mathfrak{u} \hat{=} p$, $\mathfrak{v} \hat{=} q$,
- $\mathfrak{A}(t) \hat{=} A \ \forall t$ with $D(A) = H^1(\Omega)$,
- $\mathfrak{B}(t)\mathfrak{u}(t) \hat{=} c_0 p(t) + Bp(t) \ \forall t$ whose domain is $L^2(\Omega)$,
- $\mathfrak{a}_0(t; \mathfrak{u}(t), \mathfrak{v}) \hat{=} a(p, q)$,
- $\mathfrak{a}_1(t; \mathfrak{u}(t), \mathfrak{v}) \hat{=} b(p, q)$ and $\mathfrak{a}_1'(t; \mathfrak{u}(t), \mathfrak{v}) \hat{=} b'(p, q) = \left((c_0 + B)' p, p\right)_{L^2(\Omega)} \equiv 0$ for all $p, q \in Q$ independent of t,
- $\Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}(\mathfrak{a}_1(t; \mathfrak{u}(t), \mathfrak{v})) \hat{=} \frac{\mathrm{d}}{\mathrm{d}t} b(p(t), q)$.
- The Conditions (3.81) on $\mathfrak{a}_0(t; \mathfrak{u}(t), \mathfrak{v})$ and $\mathfrak{a}_1(t; \mathfrak{u}(t), \mathfrak{v})$ are satisfied by $a(p, q)$ and $b(p, q)$ as shown above. Especially (3.81g) is satisfied by combining (3.69) and (3.73).
- As $\overline{h}(t)$ is a linear form on Q for every $t \in [0, t_{\text{end}}]$ according to [234, Example 10.7], we can find a representative $g(t)$ of $\overline{h}$ in $\left(Q, (\cdot, \cdot)_Q\right)$ according to the Riesz Representation Theorem 2.58 which we can extend to $(-\infty, 0)$ by 0.

Thus, Theorem 3.11 is applicable.

Here, the bilinear form $\mathfrak{a}_1(t;\mathfrak{u}(t),\mathfrak{v}) \hat{=} b(p,q)$ is even an inner product. Thus, we can get a better result on the regularity of p. In order to do so, we return to (3.71), represent $L^2(\Omega)$ equipped with $b(p,q)$ as inner product by $\mathscr{H}$ and apply the following theorem [234].

Theorem 3.12 *Let $\mathscr{H}$, Q be two separable Hilbert spaces such that $\mathscr{H}$ can be identified with its dual and $Q \subset \mathscr{H} \subset Q'$ with continuous and dense imbeddings. Let $a(\cdot,\cdot)$ be a continuous bilinear form which fulfills the modified coercivity condition of (3.69) and $A : Q \to Q'$ is its associated bounded linear operator. Assume that the functions $\overline{h} \in L^2((0,t_{\text{end}}),Q')$ as well as $p^0 \in \mathscr{H}$ are given and $0 < t_{\text{end}} < \infty$. Then the problem*

$$\frac{\mathrm{d}p}{\mathrm{d}t} + A(p) = \overline{h}, \quad p(0) = p^0 \tag{3.83}$$

has a unique solution $p \in L^2((0,t_{\text{end}}),Q) \cap H^1((0,t_{\text{end}}),Q')$ which continuously depends on $\overline{h}$ and p^0 and $p \in C([0,t_{\text{end}}],\mathscr{H})$.

Proof See [234, Theorem 10.3 and Lemma 10.4] as well as [292, Theorem 26.1].

Remark 3.13 $p \in C([0,t_{\text{end}}],\mathscr{H})$ is equivalent to $p \in C\big([0,t_{\text{end}}],L^2(\Omega)\big)$.

The problem with the formulation of (3.71) as in Theorem 3.12 is that it is not directly clear which conditions have to be satisfied by the initial function p^0. Therefore, we consider Formulations (3.80) and (3.82). In (3.80) we need $\mathfrak{B}\mathfrak{u}(0) = \mathfrak{B}\mathfrak{u}^0$ given in the range $R(\mathfrak{B})$ of $\mathfrak{B}$ which is basically the same as $\mathfrak{u}(0) = \mathfrak{u}^0$ given in $D(\mathfrak{B})$ and can be interpreted as either

$$\mathfrak{u}(0) = \mathfrak{u}^0 \in \mathfrak{H}, \quad \mathfrak{B}\mathfrak{u}^0 \in \mathfrak{H} \text{ exists, or} \tag{3.84a}$$

$$\left(\mathfrak{B}\mathfrak{u}^0,\mathfrak{v}\right)_{\mathfrak{H}} = \mathfrak{b}\left(\mathfrak{u}^0,\mathfrak{v}\right) \ \forall \mathfrak{v} \in \mathfrak{H} \tag{3.84b}$$

if $\mathfrak{b}(\cdot,\cdot)$ refers to the bilinear form to which $\mathfrak{B}$ is associated. If we consider (3.82), it is clear that we have to know $\mathfrak{a}_1(t;\mathfrak{u}^0,\mathfrak{v})$ which in our case means $b(p^0,q)$. This leads to several equivalent possibilities how to choose an initial condition:

(IC.1) Let $(c_0 + B)p^0 = \zeta^0$ with some $\zeta^0 \in L^2(\Omega)$ be given. Then $(c_0 p^0 + Bp^0, q)_{L^2(\Omega)}$ is defined for every $q \in L^2(\Omega)$. If the Dirichlet boundary condition on u is valid in the limit $t \to 0$, $b(\cdot,\cdot)$ defines an inner product, $p \mapsto (c_0 + B)p$ is bijective and we can determine p^0 from ζ^0. If the Dirichlet boundary condition on u does not hold in the limit, ζ^0 may be determined by choosing p^0 and $\nabla_x \cdot u^0$ [254].

(IC.2) Choose $p^0 \in L^2(\Omega)$ and let the Dirichlet boundary condition on u still be valid in the limit $t \to 0$. Then Bp^0 is well-defined and $b(p^0,q) = (c_0 p^0 + Bp^0, q)_{L^2(\Omega)}$ is defined for every $q \in L^2(\Omega)$ because $b(\cdot,\cdot)$ is an inner product on $L^2(\Omega)$. u^0 is determined by p^0 and the boundary conditions as the weak form of the Cauchy-Navier Equation (3.55) is still valid in the limit $t \to 0$.

(IC.3) Choose $u^0 = \hat{u}^0 + u^0_{\mathrm{ce}} \in V$. Then p^0 can be determined by (3.55) if $f(\cdot, 0)$ is known because (3.55) does not depend on t and, therefore, is still valid in the limit $t \to 0$. However, p^0 can only be determined uniquely up to an additive constant. This constant can be determined, if, e.g., there is a Dirichlet boundary condition on p which is still valid in the limit $t \to 0$. With p^0 known, we can then determine ζ^0 and proceed as in (IC.1).

As mentioned above, the condition $\overline{h} \in L^2((0, t_{\mathrm{end}}), Q')$ has to be satisfied by $\partial_t(\nabla_x \cdot u_{\mathrm{ce}})$, too. Because of the assumption $u_{\mathrm{ce}} \in V$, we have $\nabla_x \cdot u_{\mathrm{ce}} \in L^2(\Omega)$ for each $t \in (0, t_{\mathrm{end}})$. Thus, we demand $\partial_t(\nabla_x \cdot u_{\mathrm{ce}}) \in L^2((0, t_{\mathrm{end}}), L^2(\Omega)) \subset L^2((0, t_{\mathrm{end}}), Q')$. In order to transfer this to conditions on f and t_n we define

$$\delta_{t,\theta} u = \frac{u(x, t+\theta) - u(x, t)}{\theta} \,. \tag{3.85}$$

Evidently, one has

$$e(\delta_{t,\theta} u_{\mathrm{ce}}, v) = (\delta_{t,\theta} f, v)_{\mathcal{L}^2(\Omega)} + (\delta_{t,\theta} t_n, v)_{\mathcal{L}^2(\Gamma_t)} \quad \forall v \in V \,. \tag{3.86}$$

Considering $v = \delta_{t,\theta} u_{\mathrm{ce}}$, we get

$$|e(\delta_{t,\theta} u_{\mathrm{ce}}, \delta_{t,\theta} u_{\mathrm{ce}})| \leq \left|(\delta_{t,\theta} f, \delta_{t,\theta} u_{\mathrm{ce}})_{\mathcal{L}^2(\Omega)}\right| + \left|(\delta_{t,\theta} t_n, \delta_{t,\theta} u_{\mathrm{ce}})_{\mathcal{L}^2(\Gamma_t)}\right| \,. \tag{3.87}$$

Next we use the coercivity of $e(\cdot\,, \cdot)$ (see (3.60)) and interpret f and t_n as linear functionals on $\mathcal{H}^1(\Omega)$ and $\mathcal{H}^{\frac{1}{2}}(\Gamma_t)$, respectively, to get the estimate

$$\begin{aligned} &\|\delta_{t,\theta} u_{\mathrm{ce}}\|^2_{\mathcal{H}^1(\Omega)} \\ \leq & C \left(\|\delta_{t,\theta} f\|_{\mathcal{H}^{-1}(\Omega)} \, \|\delta_{t,\theta} u_{\mathrm{ce}}\|_{\mathcal{H}^1(\Omega)} + \|\delta_{t,\theta} t_n\|_{\mathcal{H}^{-\frac{1}{2}}(\Gamma_t)} \, \|\delta_{t,\theta} u_{\mathrm{ce}}\|_{\mathcal{H}^{\frac{1}{2}}(\Gamma_t)} \right) \,. \end{aligned} \tag{3.88}$$

Because of the continuity of the trace operator (Theorem 2.41), we can estimate $\|\delta_{t,\theta} u_{\mathrm{ce}}\|_{\mathcal{H}^{\frac{1}{2}}(\Gamma_t)}$ by $\|\delta_{t,\theta} u_{\mathrm{ce}}\|_{\mathcal{H}^1(\Omega)}$ and, thus, divide by $\|\delta_{t,\theta} u_{\mathrm{ce}}\|_{\mathcal{H}^1(\Omega)}$. Furthermore, we estimate $\|\delta_{t,\theta} u_{\mathrm{ce}}\|_{\mathcal{H}^1(\Omega)}$ on the left-hand side by $\|\nabla_x \cdot (\delta_{t,\theta} u_{\mathrm{ce}})\|_{L^2(\Omega)}$ which leads to

$$\|\nabla_x \cdot (\delta_{t,\theta} u_{\mathrm{ce}})\|_{L^2(\Omega)} \leq C \left(\|\delta_{t,\theta} f\|_{\mathcal{H}^{-1}(\Omega)} + \|\delta_{t,\theta} t_n\|_{\mathcal{H}^{-\frac{1}{2}}(\Gamma_t)} \right) \,. \tag{3.89}$$

This is still valid if we take the $L^2((0, t_{\mathrm{end}}), \cdot)$-norm on both sides. More correctly, we should not consider the whole interval $(0, t_{\mathrm{end}})$ but an interval (a, b) with $[a, b] \subset (0, t_{\mathrm{end}})$, $\theta < \min(a, t_{\mathrm{end}} - b)$.

$$\begin{aligned} &\|\nabla_x \cdot (\delta_{t,\theta} u_{\mathrm{ce}})\|_{L^2((a,b), L^2(\Omega))} \\ \leq & C \left(\|\delta_{t,\theta} f\|_{L^2((a,b), \mathcal{H}^{-1}(\Omega))} + \|\delta_{t,\theta} t_n\|_{L^2((a,b), \mathcal{H}^{-\frac{1}{2}}(\Gamma_t))} \right) \,. \end{aligned} \tag{3.90}$$

In order to proceed from the difference quotient to the weak derivative, we consider Lemma 2.45 and assume that the weak derivatives $\partial_t f$ and $\partial_t t_n$ exist in $L^2\big((0,t_{\text{end}}),\mathcal{H}^{-1}(\Omega)\big)$ and $L^2\Big((0,t_{\text{end}}),\mathcal{H}^{-\frac{1}{2}}(\Gamma_t)\Big)$, respectively. Thus, we can estimate the right-hand side to get

$$\begin{aligned}&\|\nabla_x\cdot(\delta_{t,\theta}u_{\text{ce}})\|_{L^2((a,b),L^2(\Omega))}\\ \leq& C\left(\|\partial_t f\|_{L^2((0,t_{\text{end}}),\mathcal{H}^{-1}(\Omega))}+\|\partial_t t_n\|_{L^2((0,t_{\text{end}}),\mathcal{H}^{-\frac{1}{2}}(\Gamma_t))}\right).\end{aligned}\tag{3.91}$$

Because the right-hand side is now constant for any $[a,b]\subset(0,t_{\text{end}})$ and $\theta<\min(a,t_{\text{end}}-b)$, we finally have

$$\begin{aligned}&\|\nabla_x\cdot(\partial_t u_{\text{ce}})\|_{L^2((0,t_{\text{end}}),L^2(\Omega))}\\ \leq& C\left(\|\partial_t f\|_{L^2((0,t_{\text{end}}),\mathcal{H}^{-1}(\Omega))}+\|\partial_t t_n\|_{L^2((0,t_{\text{end}}),\mathcal{H}^{-\frac{1}{2}}(\Gamma_t))}\right).\end{aligned}\tag{3.92}$$

Note that this estimate is also true if on the left-hand side we have $\|\partial_t u_{\text{ce}}\|_{L^2((0,t_{\text{end}}),\mathcal{H}^1(\Omega))}$ which we get if we do not estimate $\|\delta_{t,\theta}u_{\text{ce}}\|_{\mathcal{H}^1(\Omega)}$ by $\|\nabla_x\cdot(\delta_{t,\theta}u_{\text{ce}})\|_{L^2(\Omega)}$. A similar estimate holds for $\hat{u}$ because p belongs to $C([0,t_{\text{end}}],\mathscr{H})$.

We now summarize the results of this section.

Theorem 3.14 (Existence and Uniqueness in Quasistatic Poroelasticity) *Let $\Omega\subset\mathbb{R}^3$ be a bounded domain with Lipschitz-boundary $\partial\Omega$ and let $\Gamma_d\subset\partial\Omega$ be measurable with* $\text{meas}(\Gamma_d)>0$.

Let $\lambda>-\frac{2}{3}\mu$, $\mu>0$, $\alpha\neq 0$, $c_0\neq 0$, and $k>0$.

Let $f(\cdot,t)\in\mathcal{L}^2(\Omega)$ for every $t\in(0,t_{\text{end}})$ with $\partial_t f\in L^2((0,t_{\text{end}}),\mathcal{H}^{-1}(\Omega))$, $t_n(\cdot,t)\in\mathcal{H}^{-\frac{1}{2}}(\Gamma_t)$ for every $t\in(0,t_{\text{end}})$ with $\partial_t t_n\in L^2((0,t_{\text{end}}),\mathcal{H}^{-\frac{1}{2}}(\Gamma_t))$, $h\in L^2((0,t_{\text{end}}),L^2(\Omega))$, and $v_{f,n}\in L^2((0,t_{\text{end}}),H^{-\frac{1}{2}}(\Gamma_f))$.

Let one of the initial conditions (IC.1), (IC.2), (IC.3) *be given or another one such that $b(p(\cdot,0),q)$ is known for every $q\in Q$.*

Then there exists a unique pair (u,p) of functions which is a weak solution of the QEP with $u\in V$ $\forall t\in[0,t_{\text{end}}]$ and $\partial_t u\in L^2((0,t_{\text{end}}),\mathcal{H}^1(\Omega))$ as well as $p\in L^2((0,t_{\text{end}}),Q)\cap H^1((0,t_{\text{end}}),Q')$ and $p\in C\big([0,t_{\text{end}}],L^2(\Omega)\big)$.

Remark 3.15 The existence and uniqueness of solutions for the dynamic equations (see (3.45)) can be found in [62, 210] in the context of thermoelasticity.

The existence and uniqueness for more general equations of poroelasticity which consider non-linear effects is proven in [25–28, 242].

Without imposing definiteness conditions of material elasticities, the uniqueness of solutions can still be shown [7].

Stability results in the context of quasistatic thermoelasticity can be found in [231] and (for a finite element formulation) for dynamic poroelasticity in [245] and the references therein. However, the time derivative of the pressure is omitted in [245].

Chapter 4
Boundary Layer Potentials in Poroelasticity

As for most partial differential equations, analytic solutions to initial boundary value problems are hardly known for the quasistatic equations of poroelasticity (3.49). Therefore, approximate solutions are computed numerically. It is well known that the costs of such numerical methods highly depends on the dimension of the problem. Increasing the spatial dimension from two to three claims a much higher demand for memory and computational power. Thus, it is advantageous if the dimension can be reduced. One way to do this is the use of boundary integral equations.

In this chapter, we will derive a set of boundary integral equations for quasistatic poroelasticity. For this purpose, we consider the homogeneous quasistatic equations of poroelasticity (QEP) without a fluid force, i.e., $g = 0$, in dimensionless form

$$-\frac{\lambda+\mu}{\mu}\nabla_x\left(\nabla_x\cdot u\right) - \nabla_x^2 u + \alpha\nabla_x p = f\,, \tag{4.1a}$$

$$\partial_t\left(c_0\mu p + \alpha\left(\nabla_x\cdot u\right)\right) - \nabla_x^2 p = h\,. \tag{4.1b}$$

In the following chapters, we assume $f = 0$ and $h = 0$. Here and further on, we omit the $\check{}$ to mark dimensionless quantities if not necessary to avoid confusion. Starting from these equations, we will derive formulas similar to Green's identities as well as the fundamental solutions of these equations. By combining fundamental solutions and Green's identities, we obtain representation formulas for a solution (u, p) of (4.1) from which we can deduce layer potentials similar to those well known in potential theory for the Laplace equation or the heat equation.

The differential operator of quasistatic poroelasticity can be defined as

$$L^{\mathrm{pe}}(u,p) = \begin{pmatrix} -\frac{\lambda+\mu}{\mu}\nabla_x\left(\nabla_x\cdot u\right) - \nabla_x^2 u + \alpha\nabla_x p \\ \partial_t\left(c_0\mu p + \alpha\left(\nabla_x\cdot u\right)\right) - \nabla_x^2 p \end{pmatrix}\,. \tag{4.2}$$

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4_4

The results of this chapter were already published in parts in two articles in peer-reviewed journals [14, 17] and a book chapter [15].

4.1 Green's Identities in Poroelasticity

Let us first recall the relation of the stress tensor $\boldsymbol{\sigma}$ to the displacement vector u. In dimensionless form, this is given by

$$\boldsymbol{\sigma}_{ij}(u) = \frac{\lambda}{\mu}\frac{\partial u_k}{\partial x_k}\delta_{ij} + \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{4.3a}$$

$$= C_{ijkl}\boldsymbol{\epsilon}_{kl}(u) \tag{4.3b}$$

$$\Leftrightarrow \boldsymbol{\sigma}(u) = \frac{\lambda}{\mu}(\nabla_x \cdot u)\boldsymbol{I} + \nabla_x u + (\nabla_x u)^T \tag{4.3c}$$

$$\Rightarrow \boldsymbol{\sigma}^{\mathrm{pe}}(u,p) = \boldsymbol{\sigma}(u) - \alpha \boldsymbol{I} p \tag{4.3d}$$

with the Kronecker delta δ_{ij} and the three-dimensional second-rank unit tensor $\boldsymbol{I}$. Note that the actual values of the fourth-rank tensor C_{ijkl} used in the dimensionless formulation here differ from the ones given in (3.23) such that we now have

$$C_{ijkl}(x,t) = \frac{\lambda}{\mu}\delta_{ij}\delta_{kl} + \left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) \tag{4.4}$$

according to the non-dimensionalizing that has been used to get (3.46). The dimensionless form of Darcy's Law (3.44) in the absence of fluid forces is

$$v_f(x,t) = -\nabla_x p(x,t)\ . \tag{4.5}$$

To ensure that all of the following integrals exist, we make the rather strong assumption, that u, v are in $C^1\left((0,t_{\mathrm{end}}), \mathcal{C}^2(\Omega)\right)$ and p, q are in $C^1\big((0,t_{\mathrm{end}}), C^2(\Omega)\big)$. Starting by the weak formulation given in (3.54) and performing two integrations by parts, we get

$$\begin{aligned}
&-\int_{\Omega}\left(\partial_{y_j}\boldsymbol{\sigma}^{\mathrm{pe}}_{ij,y}(u(y,\tau),p(y,\tau))\right) v_i(y,\tau)\,\mathrm{d}V(y)\\
&=\int_{\Omega}\boldsymbol{\sigma}^{\mathrm{pe}}_{ij,y}(u(y,\tau),p(y,\tau))\left(\partial_{y_j}v_i(y,\tau)\right)\,\mathrm{d}V(y)\\
&\quad-\int_{\Gamma}\boldsymbol{\sigma}^{\mathrm{pe}}_{ij,y}(u(y,\tau),p(y,\tau))\ v_i(y,\tau)\,n_j(y)\,\mathrm{d}S(y)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\Omega} \Big(C_{ijkl} \boldsymbol{\epsilon}_{kl,y}(u(y,\tau)) \boldsymbol{\epsilon}_{ij,y}(v(y,\tau)) - \alpha p(y,\tau) \left(\partial_{y_i} v_i(y,\tau)\right) \Big) \, \mathrm{d}V(y) \\
&\quad - \int_{\Gamma} \boldsymbol{\sigma}^{\mathrm{pe}}_{ij,y}(u(y,\tau), p(y,\tau)) \; v_i(y,\tau) \, n_j(y) \, \mathrm{d}S(y) \\
&= \int_{\Omega} \Big(- u_i(y,\tau) \left(\partial_{y_j} \boldsymbol{\sigma}_{ij,y}(v(y,\tau))\right) + \alpha \left(\partial_{y_i} p(y,\tau)\right) v_i(y,\tau) \Big) \, \mathrm{d}V(y) \\
&\quad + \int_{\Gamma} \Big(u_j(y,\tau) \boldsymbol{\sigma}_{ij,y}(v(y,\tau)) \, n_i(y) - \alpha \, p(y,\tau) \, v_i(y,\tau) \, n_i(y) \Big) \, \mathrm{d}S(y) \\
&\quad - \int_{\Gamma} \boldsymbol{\sigma}^{\mathrm{pe}}_{ij,y}(u(y,\tau), p(y,\tau)) \; v_i(y,\tau) \, n_j(y) \, \mathrm{d}S(y) \, . \qquad (4.6)
\end{aligned}
$$

The subscript y, separated by a comma, is added here to state that the derivatives which have to be taken to get $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$, respectively, are taken with respect to y. This will be of importance for Green's Third Identity.

Equation (4.1b) can be treated similarly. For reasons of clarity and comprehensibility, let us first consider the Laplacian part. From this, we get

$$
\begin{aligned}
&- \int_{\Omega} \left(\partial_{y_i} \partial_{y_i} p(y,\tau)\right) q(y,\tau) \, \mathrm{d}V(y) \\
&= \int_{\Omega} \left(\partial_{y_i} p(y,\tau)\right) \left(\partial_{y_i} q(y,\tau)\right) \, \mathrm{d}V(y) - \int_{\Gamma} \left(\left(\partial_{y_i} p(y,\tau)\right) n_i(y)\right) q(y,\tau) \, \mathrm{d}S(y) \\
&= - \int_{\Omega} \left(\partial_{y_i} \partial_{y_i} q(y,\tau)\right) p(y,\tau) \, \mathrm{d}V(y) + \int_{\Gamma} p(y,\tau) \left(\left(\partial_{y_i} q(y,\tau)\right) n_i(y)\right) \, \mathrm{d}S(y) \\
&\quad - \int_{\Gamma} \left(\left(\partial_{y_i} p(y,\tau)\right) n_i(y)\right) q(y,\tau) \, \mathrm{d}S(y) \, . \qquad (4.7)
\end{aligned}
$$

For the time derivative, we also perform an integration by parts with respect to time on the interval $(0, t_{\mathrm{end}})$ to obtain

$$
\begin{aligned}
&\int_0^{t_{\mathrm{end}}} \int_{\Omega} \partial_{\tau} \left(c_0 \mu p(y,\tau) + \alpha \left(\partial_{y_i} u_i(y,\tau)\right)\right) q(y,\tau) \, \mathrm{d}V(y) \mathrm{d}\tau \\
&= - \int_0^{t_{\mathrm{end}}} \int_{\Omega} \left(c_0 \mu p(y,\tau) + \alpha \left(\partial_{y_i} u_i(y,\tau)\right)\right) \left(\partial_{\tau} q(y,\tau)\right) \, \mathrm{d}V(y) \mathrm{d}\tau
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{\Omega}\left(c_0\mu p(y,t_{\text{end}})+\alpha\left(\partial_{y_i}u_i(y,t_{\text{end}})\right)\right)q(y,t_{\text{end}})\,\mathrm{d}V(y)\\
&-\int_{\Omega}\left(c_0\mu p(y,0)+\alpha\left(\partial_{y_i}u_i(y,0)\right)\right)q(y,0)\,\mathrm{d}V(y)\\
=&-\int_0^{t_{\text{end}}}\int_{\Omega}\left(c_0\mu p(y,\tau)\right)\left(\partial_\tau q(y,\tau)\right)-\left(\alpha u_i(y,\tau)\right)\left(\partial_{y_i}\left(\partial_\tau q(y,\tau)\right)\right)\,\mathrm{d}V(y)\mathrm{d}\tau\\
&-\int_0^{t_{\text{end}}}\int_{\Gamma}\alpha\left(u_i(y,\tau)n_i(y)\right)\left(\partial_\tau q(y,\tau)\right)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&+\int_{\Omega}\left(c_0\mu p(y,t_{\text{end}})+\alpha\left(\partial_{y_i}u_i(y,t_{\text{end}})\right)\right)q(y,t_{\text{end}})\,\mathrm{d}V(y)\\
&-\int_{\Omega}\left(c_0\mu p(y,0)+\alpha\left(\partial_{y_i}u_i(y,0)\right)\right)q(y,0)\,\mathrm{d}V(y)\,. \qquad (4.8)
\end{aligned}
$$

Once again, it becomes clear that a suitable initial condition has to provide the fluid content at $t=0$, i.e., $\zeta(y,0)=c_0\mu p(y,0)+\alpha\left(\nabla_y\cdot u(y,0)\right)$.

By combining the respective parts of (4.6)–(4.8), we obtain

Theorem 4.1 (Green's First Identity in Poroelasticity) *Let* $u,v\in C^1\left((0,t_{\text{end}}),\mathcal{C}^2\left(\overline{\Omega}\right)\right)$ *and* $p,q\in C^1\left((0,t_{\text{end}}),C^2(\overline{\Omega})\right)$. *Then we have*

$$
\begin{aligned}
&\int_0^{t_{\text{end}}}\int_{\Omega}\Bigg[\left(-\frac{\lambda+\mu}{\mu}\partial_{y_i}\partial_{y_j}u_j(y,\tau)-\partial_{y_j}\partial_{y_j}u_i(y,\tau)+\alpha\partial_{y_i}p(y,\tau)\right)v_i(y,\tau)\\
&\qquad+\left(\partial_\tau\left(c_0\mu p(y,\tau)+\alpha\partial_{y_j}u_j(y,\tau)\right)\right.\\
&\qquad\left.-\partial_{y_j}\partial_{y_j}p(y,\tau)\right)q(y,\tau)\Bigg]\,\mathrm{d}V(y)\mathrm{d}\tau\\
=&\int_0^{t_{\text{end}}}\int_{\Omega}\Bigg[C_{ijkl}\boldsymbol{\epsilon}_{kl,y}(u(y,\tau))\boldsymbol{\epsilon}_{ij,y}(v(y,\tau))-\alpha p(y,\tau)\left(\partial_{y_i}v_i(y,t-\tau)\right)\\
&\qquad+\left(\partial_\tau\left(c_0\mu p(y,\tau)+\alpha\partial_{y_j}u_j(y,\tau)\right)\right)q(y,\tau)\\
&\qquad+\left(\partial_{y_j}p(y,\tau)\right)\left(\partial_{y_j}q(y,\tau)\right)\Bigg]\,\mathrm{d}V(y)\mathrm{d}\tau
\end{aligned}
$$

$$-\int_0^{t_{\text{end}}}\int_\Gamma \Big[\boldsymbol{\sigma}_{ij,y}(u(y,\tau))n_j(y)v_i(y,\tau) - \alpha\left(p(y,\tau)n_i(y)\right)v_i(y,\tau) + \left(\left(\partial_{y_j}p(y,\tau)\right)n_j(y)\right)q(y,\tau)\Big]\,\mathrm{d}S(y)\mathrm{d}\tau\,. \tag{4.9}$$

Remark 4.2 The term $\boldsymbol{\sigma}_{ij,y}(u(y,\tau))n_j(y)$ is known in classical linear elasticity theory for the Cauchy-Navier equation as its co-normal derivative operator $\mathrm{T}_y^{\mathrm{ce}}(u(y))$. It can also be written as (see, e.g., [91, Chapter 4])

$$\begin{aligned}\mathrm{T}_y^{\mathrm{ce}}(u(y)) =& \boldsymbol{\sigma}(u(y))n(y)\\ =& \frac{\lambda}{\mu}\left(\nabla_y\cdot u(y)\right)n(y) + 2\left(\nabla_y u(y)\right)n(y) + n(y)\wedge\left(\nabla_y\wedge u(y)\right)\,.\end{aligned} \tag{4.10}$$

Theorem 4.3 (Green's Second Identity in Poroelasticity) *Let* $u, v \in C^1\left([0,t_{\text{end}}], \mathcal{C}^2\left(\overline{\Omega}\right)\right)$ *and* $p, q \in C^1\left([0,t_{\text{end}}], C^2(\overline{\Omega})\right)$. *Then we have*

$$\begin{aligned}&\int_0^{t_{\text{end}}}\int_\Omega \Bigg(\Big[-\frac{\lambda+\mu}{\mu}\partial_{y_i}\partial_{y_j}v_j(y,\tau) - \partial_{y_j}\partial_{y_j}v_i(y,\tau) + \alpha\partial_{y_i}\partial_\tau q(y,\tau)\Big]u_i(y,\tau)\\ &\qquad + \Big[-c_0\mu\partial_\tau q(y,\tau) - \alpha\partial_{y_j}v_j(y,\tau) - \partial_{y_j}\partial_{y_j}q(y,\tau)\Big]p(y,\tau)\\ &\qquad - \Bigg\{v_i(y,\tau)\Big[-\frac{\lambda+\mu}{\mu}\partial_{y_i}\partial_{y_j}u_j(y,\tau) - \partial_{y_j}\partial_{y_j}u_i(y,\tau) + \alpha\partial_{y_i}p(y,\tau)\Big]\\ &\qquad + q(y,\tau)\Big[\partial_\tau\left(c_0\mu p(y,\tau) + \alpha\partial_{y_j}u_j(y,\tau)\right) - \partial_{y_j}\partial_{y_j}p(y,\tau)\Big]\Bigg\}\Bigg)\,\mathrm{d}V(y)\mathrm{d}\tau\\ &+ \int_\Omega q(y,t_{\text{end}})\left(c_0\mu p(y,t_{\text{end}}) + \alpha\partial_{y_i}u_i(y,t_{\text{end}})\right)\,\mathrm{d}V(y)\\ =&\int_0^{t_{\text{end}}}\int_\Gamma \Big(v_i(y,\tau)\Big[\boldsymbol{\sigma}_{ij,y}(u(y,\tau))n_j(y) - \alpha p(y,\tau)n_i(y)\Big]\\ &\qquad + q(y,\tau)\Big[\left(\partial_{y_j}p(y,\tau)\right)n_j(y)\Big]\Big)\,\mathrm{d}S(y)\mathrm{d}\tau\end{aligned}$$

$$
\begin{aligned}
-\int_0^{t_{\text{end}}}\int_\Gamma \Bigg(\Big[\boldsymbol{\sigma}_{ij,y}(v(y,\tau))n_j(y) - \alpha\partial_\tau q(y,\tau)n_i(y) \Big] u_i(y,\tau) \\
+ \Big[\left(\partial_{y_j} q(y,\tau)\right) n_j(y) \Big] p(y,\tau) \Bigg) \, \mathrm{d}S(y)\mathrm{d}\tau \\
+ \int_\Omega q(y,0) \left(c_0 \mu p(y,0) + \alpha \partial_{y_i} u_i(y,0) \right) \, \mathrm{d}V(y) \, . \qquad (4.11)
\end{aligned}
$$

From Green's Second Identity, we can identify the adjoint equations of (4.1) to be

$$
-\frac{\lambda+\mu}{\mu}\nabla_y \left(\nabla_y \cdot v(y,\tau) \right) - \nabla_y^2 v(y,\tau) + \alpha \nabla_y \left(\partial_\tau q(y,\tau) \right) = 0 \, , \qquad (4.12a)
$$

$$
-\partial_\tau \left(c_0 \mu q(y,\tau) \right) - \alpha \left(\nabla_y \cdot v(y,\tau) \right) - \nabla_y^2 q(y,\tau) = 0 \, . \qquad (4.12b)
$$

Thus, the adjoint operator of the differential operator of quasistatic poroelasticity is

$$
(L^{\text{pe}})^* (v,q) = \begin{pmatrix} -\frac{\lambda+\mu}{\mu}\nabla_y \left(\nabla_y \cdot v \right) - \nabla_y^2 v + \alpha \nabla_y \left(\partial_\tau q \right) \\ -\partial_\tau \left(c_0 \mu q \right) - \alpha \left(\nabla_y \cdot v \right) - \nabla_y^2 q \end{pmatrix} \, . \qquad (4.13)
$$

Remark 4.4 It is of course possible to reduce the regularity requirements on u, v, p, and q by choosing appropriate Sobolev spaces. The reader is referred to an article by Costabel [57] for an overview how this can be done in the context of the heat equation. Please note that as we deal with more complicated differential operators in quasistatic poroelasticity, the restrictions on the Sobolev spaces have to be adapted, especially with regard to the terms in which derivatives with respect to time and with respect to space are mixed.

4.2 Fundamental Solutions

According to the Encyclopedia of Mathematics [128], a fundamental solution is

> [a] solution of a partial differential equation $Lu(x) = 0$, $x \in \mathbb{R}^n$, with coefficients of class C^∞, in the form of a function $l(x,y)$ that satisfies, for fixed $y \in \mathbb{R}^n$, the equation
>
> $$
> L\, l(x,y) = \delta(x-y), \; x \neq y \, , \qquad (4.14)
> $$
>
> which is interpreted in the sense of the theory of generalized functions.

Moreover,

> [t]he Green function of a boundary value problem for a linear differential equation is [understood to be] the fundamental solution of this equation satisfying homogeneous boundary conditions.

Thus, fundamental solutions are independent of the domain and boundary conditions while a Green function may vary with the domain on which a differential equation is studied as well as with the associated boundary conditions. Within this thesis, we will use fundamental solutions in the sense as given above.

Differential equations with an explicit dependency on time are usually considered on domains that can be written as $\overline{\Omega} \times [0, t_{\text{end}}]$, where $\Omega \subset \mathbb{R}^n$ is an open domain. With the notation of Sect. 2.5, we write

$$\begin{aligned} Lu(x,t) = &\sum_{i,j=1}^{n} a_{ij}(x,t)\partial_{x_i}\partial_{x_j}u(x,t) + \sum_{i=1}^{n} b_i(x,t)\partial_{x_i}u(x,t) \\ &+ c(x,t)u(x,t) - \partial_t u(x,t) \ . \end{aligned} \tag{4.15}$$

and make the following assumptions:

(i) L is uniformly parabolic in $\mathbb{R}^n \times [0, t_{\text{end}}]$ as defined in Definition 2.54,
(ii) The coefficients of L are bounded continuous functions on $\mathbb{R}^n \times [0, t_{\text{end}}]$ and for all $(x,t) \in \mathbb{R}^n \times [0, t_{\text{end}}]$, $(y,\tau) \in \mathbb{R}^n \times [0, t_{\text{end}}]$

$$\left|a_{ij}(x,t) - a_{ij}(y,\tau)\right| \leq A\left(\|x-y\|^{\beta} + |t-\tau|^{\frac{\beta}{2}}\right) \ , \tag{4.16a}$$

$$|b_i(x,t) - b_i(y,t)| \leq A\,\|x-y\|^{\beta} \ , \tag{4.16b}$$

$$|c(x,t) - c(y,t)| \leq A\,\|x-y\|^{\beta} \ , \tag{4.16c}$$

for constants $A > 0$ and $0 < \beta < 1$.

Hence, there are positive constants λ_0, λ_1 such that

$$\lambda_0 \|x-y\|^2 \leq \sum_{i,j=1}^{n} a^{ij}(\xi,\tau)\,(x_i - y_i)\left(x_j - y_j\right) \leq \lambda_1 \|x-y\|^2 \tag{4.17}$$

with $\tau \in [0, t_{\text{end}}]$ and $(a^{ij}(x,t))$ the inverse matrix to $(a_{ij}(x,t))$ [101].

Under the above assumptions, Friedman [101, Section 1.1.] gives the following definition of a fundamental solution, which we adapted to our notation.

Definition 4.5 Let $\mathscr{M}$ be the set of all tuple $(x,t;y,\tau)$ with $(x,t) \in \mathbb{R}^n \times [0, t_{\text{end}}]$, $(y,\tau) \in \mathbb{R}^n \times [0, t_{\text{end}}]$, $t > \tau$. A fundamental solution of $Lu = 0$ in $\mathbb{R}^n \times [0, t_{\text{end}}]$ is a function $G : \mathscr{M} \to \mathbb{R}^m$, $(x,t;y,\tau) \mapsto G(x,t;y,\tau)$, $n, m \in \mathbb{N}$, with m being the dimension of the system of differential equations, which satisfies the conditions:

(i) For fixed (y,τ) it satisfies, as a function of (x,t) with $x \in \mathbb{R}^n$, $\tau < t < t_{\text{end}}$, the equation $LG(\cdot,\cdot;y,\tau) = 0$;

(ii) For every continuous function $f : \mathbb{R}^n \to \mathbb{R}^m$ with $\|f(x)\| \leq c \exp(h \|x\|^2)$ for all $x \in \mathbb{R}^n$ and some positive constants $c, h \in \mathbb{R}^+$, we have

$$\lim_{\substack{t \to \tau \\ t > \tau}} \int_{\mathbb{R}^n} G(x, t; y, \tau) f(y) \, \mathrm{d}y = f(x) \tag{4.18}$$

if $4h(t - \tau) < \lambda_0$ with λ_0 as in (4.17).

Remark 4.6 The second condition of the above definition is equivalent to $G(x, t; y, \tau)$ satisfying

$$\lim_{t \to \tau} G(x, t; y, \tau) = \delta(x - y) \tag{4.19}$$

in the sense of distributions which can be interpreted as an initial condition to the equation $LG(\cdot, \cdot; y, \tau) = 0$. Thus, Definition 4.5 can be seen to be in accordance with the definition given in the Encyclopedia of Mathematics [128] when interpreted on $\mathbb{R}^n \times [0, t_{\text{end}}]$ including a slightly different meaning of the Dirac delta distribution with respect to time. This becomes obvious if said definition is written in integral form as

$$L(l * f) = f,$$

for any infinitely differentiable function f with compact support in $\mathbb{R}^n$.

The above definitions can also be formulated for systems of differential equations.

Unfortunately, System (4.1) is not uniformly parabolic. On the one hand, (4.1a) has no explicit dependency on time and does not involve a time derivative of u. On the other hand, (4.1b) contains the term $\partial_t(\nabla_x \cdot u)$ which mixes derivatives with respect to time and spatial variables. As a consequence, we cannot assume that theorems for uniformly parabolic operators still hold for quasistatic poroelasticity.

The existence of fundamental solutions in the sense of distributions for differential operators with constant coefficients was established by Malgrange [184] and Ehrenpreis [69]. For a proof of this, see, e.g., [238, Theorem 8.5].

In this section, we introduce a method to calculate fundamental solutions of (3.49) based on a scheme formulated by Cheng and Detournay (see [52] and the references therein). In order to do this, it is convenient to use the dimensionless form of the QEP (4.1) and rearrange them to be expressed in terms of the displacement u and the volumetric fluid content change ζ as

$$-\frac{c_0(\lambda + \mu) + \alpha^2}{c_0 \mu} \nabla_x (\nabla_x \cdot u) - \nabla_x^2 u = f - \frac{\alpha}{c_0 \mu} \nabla_x \zeta \,, \tag{4.20a}$$

$$\partial_t \zeta - \frac{\lambda + 2\mu}{c_0 \mu (\lambda + 2\mu) + \mu \alpha^2} \nabla_x^2 \zeta = \frac{\alpha}{c_0 (\lambda + 2\mu) + \alpha^2} \nabla_x \cdot f + h \,. \tag{4.20b}$$

Here, we assumed all parameters to be scalar constants. The dimensionless fluid content ζ is given by

$$\zeta = c_0 \mu p + \alpha \left(\nabla_x \cdot u \right) \ . \tag{4.21}$$

As suggested by Biot [37], we define a scalar potential Ψ and an auxiliary displacement vector u_{ce} by

$$\zeta = \nabla_x^2 \Psi, \tag{4.22a}$$

$$u = u_{\mathrm{ce}} + \frac{\alpha}{c_0 \left(\lambda + 2\mu \right) + \alpha^2} \nabla_x \Psi \ , \tag{4.22b}$$

to get

$$-\frac{c_0 \left(\lambda + \mu \right) + \alpha^2}{c_0 \mu} \nabla_x \left(\nabla_x \cdot u_{\mathrm{ce}} \right) - \nabla_x^2 u_{\mathrm{ce}} = f \ , \tag{4.23a}$$

$$\partial_t \Psi - \frac{\lambda + 2\mu}{c_0 \mu \left(\lambda + 2\mu \right) + \mu \alpha^2} \nabla_x^2 \Psi = \frac{\alpha}{c_0 \left(\lambda + 2\mu \right) + \alpha^2} \Upsilon_1 + \Upsilon_2 \ . \tag{4.23b}$$

As pointed out in [52], an arbitrary harmonic function may be added to Ψ without violating (4.23b). In order to determine if such a harmonic function should be added to Ψ, we have to examine whether the fundamental solutions which we derive here satisfy the conditions of Definition 4.5. For the three-dimensional case, it turns out that this harmonic function is identical to zero.

The auxiliary functions Υ_1 and Υ_2 are defined via

$$\nabla_x^2 \Upsilon_1 = \nabla_x \cdot f \ , \tag{4.24a}$$

$$\nabla_x^2 \Upsilon_2 = h \ . \tag{4.24b}$$

The pore pressure p can be calculated by

$$p = \partial_t \Psi - \Upsilon_2 \tag{4.25a}$$

$$= \frac{\lambda + 2\mu}{c_0 \mu \left(\lambda + 2\mu \right) + \mu \alpha^2} \zeta + \frac{\alpha}{c_0 \left(\lambda + 2\mu \right) + \alpha^2} \Upsilon_1 \ . \tag{4.25b}$$

A list of fundamental solutions in a homogeneous isotropic poroelastic medium is given by Cheng and Detournay in [52]. However, they did not use the dimensionless form of the equations. Unfortunately, they also did not include factors in front of the source terms to point out that a certain kind of unit has to be associated with every source term. For example, if we choose, in a dimensioned setting, $h(x,t) = \delta(x) H(t)$ with Dirac's delta distribution $\delta(\cdot)$ and the Heaviside function $H(\cdot)$, the unit of h is m^{-3}although it should be s^{-1}. Thus, a more exact notation would be $h(x,t) = h_0 \delta(x) H(t)$, where the unit of h_0 is $\mathrm{m}^3\mathrm{s}^{-1}$. Note, that the unit of the factor depends on the function assigned to the right-hand side, as for $h(x,t) = h_0 \delta(x) \delta(t)$ the unit of h_0 is m^3.

Before we establish the fundamental solutions of (4.1), let us define the abbreviations

$$C_1 = \frac{\alpha}{c_0(\lambda + 2\mu) + \alpha^2} \quad \text{and} \quad C_2 = \frac{\lambda + 2\mu}{c_0\mu(\lambda + 2\mu) + \mu\alpha^2} \,. \tag{4.26}$$

We start by assuming $f \equiv 0$ and $h(x,t) = \delta(x)\delta(t)$. The corresponding fundamental solutions will be labeled with a superscript "Si" for an instantaneous (fluid) source. From (4.20b) it is clear, that ζ^{Si} is the fundamental solution of the heat equation, or heat kernel, which we will denote by

$$G^{\mathrm{Heat}}(x,t) = \frac{1}{\sqrt{\pi}^3} \frac{1}{\sqrt{4C_2 t}^3} \exp\left(-\frac{\|x\|^2}{4C_2 t}\right) \,. \tag{4.27}$$

From (4.25b), it results that

$$p^{\mathrm{Si}}(x,t) = C_2 G^{\mathrm{Heat}}(x,t) \tag{4.28}$$

because $\Upsilon_1^{\mathrm{Si}} \equiv 0$ for $f \equiv 0$. With (4.25a), we get

$$\begin{aligned} \partial_t \Psi^{\mathrm{Si}}(x,t) =& C_2 G^{\mathrm{Heat}}(x,t) + \Upsilon_2^{\mathrm{Si}} \\ \Rightarrow \Psi^{\mathrm{Si}}(x,t) =& \int_0^t C_2 G^{\mathrm{Heat}}(x,\tau) + G^{\mathrm{Harm}}(x)\delta(\tau)\,\mathrm{d}\tau \end{aligned} \tag{4.29}$$

because according to (4.24b), Υ_2^{Si} is the fundamental solution of the Laplace equation, denoted by

$$G^{\mathrm{Harm}}(x) = -\frac{1}{4\pi\,\|x\|}, \tag{4.30}$$

multiplied by the delta distribution $\delta(t)$. With (4.22b), this yields

$$u^{\mathrm{Si}}(x,t) = C_1 \left(\nabla_x \int_0^t \left(C_2 G^{\mathrm{Heat}}(x,\tau) + G^{\mathrm{Harm}}(x)\delta(\tau) \right) \mathrm{d}\tau \right). \tag{4.31}$$

Now let us consider the case $h \equiv 0$ and $f(x,t)$ proportional to $\delta(x)\delta(t)$. Because f is a vector-valued function, we can assign this product of delta distributions to each of its components. It is convenient to cover these three cases in one by assigning to f a tensor-valued function such that $f(x,t) = -\boldsymbol{I}\delta(x)\delta(t)$. The corresponding fundamental solutions will be labeled with a superscript "Fi" for an instantaneous (body) force. According to (4.23a), we obtain as a part of $\boldsymbol{u}^{\mathrm{Fi}}$ the fundamental solution of the Cauchy-Navier equation, which we denote by $\boldsymbol{u}^{\mathrm{CN}}(x)$. Further on, we

observe that $\nabla_x \cdot (\boldsymbol{I}\delta(x)\delta(t)) = \nabla_x \delta(x)\delta(t)$. Thus, $\Upsilon_1^{\mathrm{Fi}} = \nabla_x \Upsilon_2^{\mathrm{Si}}$, $\Psi^{\mathrm{Fi}} = C_1 \nabla_x \Psi^{\mathrm{Si}}$, and with (4.25a) as well as (4.22b), respectively, this yields

$$p^{\mathrm{Fi}}(x,t) = C_1 \left(\nabla_x C_2 G^{\mathrm{Heat}}(x,t) + \nabla_x G^{\mathrm{Harm}}(x)\delta(t) \right) , \tag{4.32}$$

$$\boldsymbol{u}^{\mathrm{Fi}}(x,t) = C_1^2 \nabla_x \left(\nabla_x \int_0^t C_2 G^{\mathrm{Heat}}(x,\tau) + G^{\mathrm{Harm}}(x)\delta(\tau)\, \mathrm{d}\tau \right) + \boldsymbol{u}^{\mathrm{CN}}(x)\delta(t) . \tag{4.33}$$

The formulas given so far for the fundamental solutions of the QEP clearly show that they are closely related to the well known fundamental solutions of the Laplace, heat, and Cauchy-Navier equations. There is also a relation to the system consisting of the Stokes equations, because the fundamental solution for the pressure in the Stokes equations is

$$p^{\mathrm{St}}(x) = \nabla_x G^{\mathrm{Harm}}(x) . \tag{4.34}$$

With the abbreviations

$$C_3 = \frac{c_0(\lambda + 3\mu) + \alpha^2}{2\left(c_0(\lambda + 2\mu) + \alpha^2\right)} \quad \text{and} \quad C_4 = \frac{c_0(\lambda + \mu) + \alpha^2}{c_0(\lambda + 3\mu) + \alpha^2} \tag{4.35}$$

the fundamental solutions can be explicitly expressed as

$$p^{\mathrm{Si}}(x,t) = C_2 \frac{1}{\sqrt{\pi}^3} \frac{1}{\sqrt{4C_2 t}^3} \exp\left(-\frac{\|x\|^2}{4C_2 t} \right) , \tag{4.36a}$$

$$u^{\mathrm{Si}}(x,t) = C_1 \frac{x}{4\pi \|x\|^3} \left(\operatorname{erf}\left(\frac{\|x\|}{\sqrt{4C_2 t}} \right) - \frac{2}{\sqrt{\pi}} \frac{\|x\|}{\sqrt{4C_2 t}} \exp\left(-\frac{\|x\|^2}{4C_2 t} \right) \right) , \tag{4.36b}$$

$$p^{\mathrm{Fi}}(x,t) = C_1 \frac{x}{4\pi \|x\|^3} \delta(t) - C_1 C_2 \frac{2}{\sqrt{\pi}^3} \frac{x}{\sqrt{4C_2 t}^5} \exp\left(-\frac{\|x\|^2}{4C_2 t} \right) , \tag{4.36c}$$

$$\begin{aligned} \boldsymbol{u}_{ki}^{\mathrm{Fi}}(x,t) = {} & C_3 \frac{1}{4\pi \|x\|} \left(\delta_{ki} + C_4 \frac{x_i x_k}{\|x\|^2} \right) \delta(t) \\ & + C_1^2 \frac{1}{4\pi \|x\|^3} \left[\left(\delta_{ik} - \frac{3 x_i x_k}{\|x\|^2} \right) \right. \\ & \times \left(\operatorname{erf}\left(\frac{\|x\|}{\sqrt{4C_2 t}} \right) - \frac{2}{\sqrt{\pi}} \frac{\|x\|}{\sqrt{4C_2 t}} \exp\left(-\frac{\|x\|^2}{4C_2 t} \right) \right) \\ & \left. + \frac{4}{\sqrt{\pi}} \frac{\|x\|}{\sqrt{4C_2 t}^3} x_i\, x_k \exp\left(-\frac{\|x\|^2}{4C_2 t} \right) \right] , \end{aligned} \tag{4.36d}$$

where $\operatorname{erf}(\cdot)$ is the Gauß error function defined by

$$\operatorname{erf}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\eta \exp(-\vartheta^2)\, \mathrm{d}\vartheta \ . \tag{4.37}$$

We summarize these fundamental solutions by

$$\boldsymbol{G}(x-y,t-\tau) = \begin{pmatrix} \boldsymbol{u}^{\mathrm{Fi}}(x-y,t-\tau) & u^{\mathrm{Si}}(x-y,t-\tau) \\ \big(p^{\mathrm{Fi}}(x-y,t-\tau)\big)^T & p^{\mathrm{Si}}(x-y,t-\tau) \end{pmatrix} \ . \tag{4.38}$$

The following definition will also come in handy:

$$\boldsymbol{u}^{\mathrm{fi}}(x-y,t-\tau) = \boldsymbol{u}^{\mathrm{Fi}}(x-y,t-\tau) - \boldsymbol{u}^{\mathrm{CN}}(x-y)\delta(t-\tau) \ , \tag{4.39a}$$

$$p^{\mathrm{fi}}(x-y,t-\tau) = p^{\mathrm{Fi}}(x-y,t-\tau) - p^{\mathrm{St}}(x-y)\delta(t-\tau) \ . \tag{4.39b}$$

Please note that we use the superscript "fi" only as defined above and not in the sense in which it is used in [65] for some slightly different solutions. For the corresponding stresses and flow vectors, see Appendix A.

Remark 4.7 On the plane $\mathbb{R}^2$, the fundamental solutions read

$$p^{\mathrm{Si}}(x,t) = \frac{1}{4\pi C_2 t} \exp\left(-\frac{\|x\|^2}{4C_2 t}\right) \ , \tag{4.40a}$$

$$u^{\mathrm{Si}}(x,t) = C_1 \frac{x}{2\pi \|x\|^2} \left(1 - \exp\left(-\frac{\|x\|^2}{4C_2 t}\right)\right) \ , \tag{4.40b}$$

$$p^{\mathrm{Fi}}(x,t) = C_1 \frac{x}{2\pi \|x\|^2} \delta(t) - C_1 C_2 \frac{2}{\pi} \frac{x}{(4C_2 t)^2} \exp\left(-\frac{\|x\|^2}{4C_2 t}\right) \ , \tag{4.40c}$$

$$\begin{aligned} \boldsymbol{u}^{\mathrm{Fi}}_{ki}(x,t) = {} & C_3 \frac{1}{2\pi} \left(-\delta_{ki} \ln(\|x\|) + C_4 \frac{x_i x_k}{\|x\|^2}\right) \delta(t) \\ & + C_1^2 \frac{1}{2\pi \|x\|^2} \left[\left(\delta_{ik} - \frac{2x_i x_k}{\|x\|^2}\right) \left(1 - \exp\left(-\frac{\|x\|^2}{4C_2 t}\right)\right) \right. \\ & \left. + \frac{2}{4C_2 t} x_i x_k \exp\left(-\frac{\|x\|^2}{4C_2 t}\right) \right] \ . \end{aligned} \tag{4.40d}$$

Please note that in the derivation of these fundamental solutions, a non-vanishing harmonic function has to be added to Ψ in order to obtain the factor $\left(1 - \exp\left(-\frac{\|x\|^2}{4C_2 t}\right)\right)$ instead of just the negative exponential function.

The main results of this thesis, which are shown for the three-dimensional case, also hold in two dimensions.

As usual, all fundamental solutions are interpreted in the sense of distributions for $x \in \mathbb{R}^3$ and $t \in [0, \infty)$. For $t > 0$, the following lemma holds.

Lemma 4.8 *For any fixed $t > 0$, the fundamental solutions $p^{\mathrm{Si}}(\cdot, t)$, $u^{\mathrm{Si}}(\cdot, t)$, $p^{\mathrm{Fi}}(\cdot, t)$, and $\boldsymbol{u}^{\mathrm{Fi}}(\cdot, t)$ are real analytic. Moreover,*

$$p^{\mathrm{Si}}(0,t) = \frac{C_2}{\sqrt{4\pi C_2 t}^3}\,, \tag{4.41a}$$

$$u^{\mathrm{Si}}(0,t) = 0\,, \tag{4.41b}$$

$$p^{\mathrm{Fi}}(0,t) = 0\,, \tag{4.41c}$$

$$\boldsymbol{u}^{\mathrm{Fi}}_{ki}(0,t) = \frac{C_1^2}{3\sqrt{4\pi C_2 t}^3}\delta_{ki}, \quad i,k \in \{1,2,3\}\,. \tag{4.41d}$$

Proof For p^{Si}, the lemma is true because of the properties of the exponential function.

For u^{Si}, we use the series expansions of $\exp(\eta)$ and $\operatorname{erf}(\eta)$ to obtain

$$\begin{aligned}
u^{\mathrm{Si}}(x,t) =& C_1 \frac{x}{4\pi \|x\|^3} \left(\operatorname{erf}\left(\frac{\|x\|}{\sqrt{4C_2 t}} \right) - \frac{2}{\sqrt{\pi}} \frac{\|x\|}{\sqrt{4C_2 t}} \exp\left(-\frac{\|x\|^2}{4C_2 t} \right) \right) \\
=& C_1 \frac{x}{4\pi \|x\|^3} \left(\frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(2n+1)} \left(\frac{\|x\|}{\sqrt{4C_2 t}} \right)^{2n+1} \right. \\
& \left. - \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\frac{\|x\|}{\sqrt{4C_2 t}} \right)^{2n+1} \right) \\
=& C_1 \frac{1}{2\sqrt{\pi}^3} \frac{x}{\|x\|^3} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(\frac{1}{2n+1} - 1 \right) \left(\frac{\|x\|}{\sqrt{4C_2 t}} \right)^{2n+1} \\
=& C_1 \frac{1}{2\sqrt{\pi}^3} \frac{x}{\|x\|^3} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!} \frac{2n}{2n+1} \left(\frac{\|x\|}{\sqrt{4C_2 t}} \right)^{2n+1} \\
=& C_1 \frac{1}{2\sqrt{\pi}^3} \frac{x}{\|x\|^3} \sum_{n=0}^{\infty} \frac{(-1)^{n+2}}{(n+1)!} \frac{2n+2}{2n+3} \left(\frac{\|x\|}{\sqrt{4C_2 t}} \right)^{2n+3} \\
=& C_1 \frac{1}{2\sqrt{\pi}^3} \frac{x}{\sqrt{4C_2 t}^3} \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} \frac{2n+2}{2n+3} \left(\frac{\|x\|}{\sqrt{4C_2 t}} \right)^{2n}\,.
\end{aligned} \tag{4.42}$$

For $t > 0$ fixed, $\sqrt{4C_2t}^3$ is a constant. The last series converges for all $x \in \mathbb{R}^3$, thus, the lemma is proven for u^{Si} as x_i, $i \in \{1, 2, 3\}$, appears as a factor in front of the series in each component of $u^{\mathrm{Si}}(x, t)$.

As we require $t > 0$, the lemma is obviously true for p^{Fi}.

For $\boldsymbol{u}^{\mathrm{Fi}}$, we use the fact that $\boldsymbol{u}^{\mathrm{Fi}}_{ki}(x, t) = \boldsymbol{u}^{\mathrm{CN}}_{ki}(x)\delta(t) + C_1 \partial_{x_k} u^{\mathrm{Si}}_i(x, t)$. Therefore, $\boldsymbol{u}^{\mathrm{Fi}}_{ki}(\cdot, t)$ is real analytic for $t > 0$ and can be written as

$$\boldsymbol{u}^{\mathrm{Fi}}_{ki}(x, t) = C_1^2 \, \frac{1}{2\sqrt{4\pi C_2 t}^3} \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} \, \frac{2n+2}{2n+3} \left(\delta_{ki} + 2n \frac{x_k x_i}{\|x\|^2} \right) \left(\frac{\|x\|}{\sqrt{4C_2 t}} \right)^{2n} . \tag{4.43}$$

For $x = 0$, only the constant term of the series ($n = 0$) remains for the diagonal elements, whereas for the non-diagonal elements, the series vanishes.

Remark 4.9 It is possible, in some sense, to ask what happens with the fundamental solutions as t approaches zero. For G^{Heat}, we have

$$\lim_{t \to 0} G^{\mathrm{Heat}}(x, t) = \delta(x) \quad \Longrightarrow \quad \lim_{t \to 0} p^{\mathrm{Si}}(x, t) = C_2 \delta(x) \tag{4.44}$$

in the sense of a weak limit for distributions. Moreover, it is easy to see that

$$\lim_{t \to 0} u^{\mathrm{Si}}(x, t) = C_1 \frac{x}{4\pi \, \|x\|^3} , \tag{4.45}$$

for $x \neq 0$. For $x = 0$, the limit depends on whether we first take the limit with respect to t and then with respect to x or vice versa. If we set $x = 0$ and then take the limit $t \to 0$, we obtain from (4.41b) that

$$\lim_{t \to 0} u^{\mathrm{Si}}(0, t) = 0 , \tag{4.46}$$

if we consider the limit as a pointwise limit.

Using the relations between p^{fi} and p^{Si} as well as between $\boldsymbol{u}^{\mathrm{fi}}$ and u^{Si}, respectively, we also obtain

$$\lim_{t \to 0} p^{\mathrm{fi}}(x, t) = C_1 \nabla_x \delta(x) , \tag{4.47}$$

$$\lim_{t \to 0} \boldsymbol{u}^{\mathrm{fi}}_{ik}(x, t) = C_1^2 \frac{1}{4\pi \, \|x\|^3} \left(\delta_{ik} - \frac{3 x_i x_k}{\|x\|^2} \right) . \tag{4.48}$$

Remark 4.10 Alternatively to the aforementioned procedure, a Laplace transformation with respect to time can be applied first and then fundamental solutions can be calculated in the Laplace domain, a method often preferred by engineers. The

convolution with respect to time occurring in the integrals above is then transformed into a simple product, i.e.,

$$\int_0^t F_1(t-\tau)F_2(\tau)\mathrm{d}\tau \longrightarrow \mathring{F}_1(s)\mathring{F}_2(s) \tag{4.49}$$

with the Laplace transformation

$$\mathring{F}(s) = \int_0^\infty F(t)\exp(-st)\mathrm{d}t\ . \tag{4.50}$$

Here, $s \in \mathbb{C}$ is the transformation parameter. See, e.g., [229, 230, 289] for details on Laplace transformation and tables of Laplace transformed functions.

The Laplace transformed fundamental solutions for $x \in \mathbb{R}^3$ are

$$\mathring{p}^{\mathrm{Si}}(x,s) = \frac{1}{4\pi\,\|x\|}\ \exp\left(-\|x\|\sqrt{\frac{s}{C_2}}\right)\ , \tag{4.51a}$$

$$\begin{aligned}\mathring{u}^{\mathrm{Si}}(x,s) =& C_1\ \frac{1}{s}\ \frac{x}{4\pi\,\|x\|^3}\left(1-\exp\left(-\|x\|\sqrt{\frac{s}{C_2}}\right)\right)\\ &- C_1\ \frac{1}{\sqrt{C_2 s}}\ \frac{x}{4\pi\,\|x\|^2}\ \exp\left(-\|x\|\sqrt{\frac{s}{C_2}}\right)\ ,\end{aligned} \tag{4.51b}$$

$$\begin{aligned}\mathring{p}^{\mathrm{Fi}}(x,s) =& C_1\ \frac{x}{4\pi\,\|x\|^3}\left(1-\exp\left(-\|x\|\sqrt{\frac{s}{C_2}}\right)\right)\\ &- C_1\ \frac{x}{4\pi\,\|x\|^2}\ \sqrt{\frac{s}{C_2}}\ \exp\left(-\|x\|\sqrt{\frac{s}{C_2}}\right)\ ,\end{aligned} \tag{4.51c}$$

$$\begin{aligned}\mathring{\boldsymbol{u}}^{\mathrm{Fi}}_{ki}(x,s) =& C_3\ \frac{1}{4\pi\,\|x\|}\left(\delta_{ki}+C_4\frac{x_i x_k}{\|x\|^2}\right)\\ &+ C_1^2\ \frac{1}{s}\ \frac{1}{4\pi\,\|x\|^3}\left(\delta_{ik}-3\frac{x_i x_k}{\|x\|^2}\right)\\ &\times\left(1-\exp\left(-\|x\|\sqrt{\frac{s}{C_2}}\right)-\frac{2}{3}\sqrt{\frac{s}{C_2}}\,\|x\|\exp\left(-\|x\|\sqrt{\frac{s}{C_2}}\right)\right)\\ &+ \frac{C_1^2}{C_2}\ \frac{1}{4\pi}\ \frac{x_i x_k}{\|x\|^3}\ \exp\left(-\|x\|\sqrt{\frac{s}{C_2}}\right)\ .\end{aligned} \tag{4.51d}$$

From these, we can directly conclude that for $s \neq 0$ all of the above Laplace transformed fundamental solutions converge to 0 as $\|x\|$ grows to infinity as well

as that they behave like $\frac{1}{\|x\|}$ as $\|x\|$ approaches 0. The latter is a difference from the behavior of the fundamental solutions in the time domain.

Using Laplace transformation, it is also possible to calculate fundamental solutions for the fully dynamic Eqs. (3.45) [48, 49, 156, 186], even when using a more complicated equation than Darcy's law to describe the fluid flow [248]. In the latter case, there exists no closed form for the dynamic fundamental solutions in the time domain. For a boundary integral formulation in case of dynamic poroelasticity, see, e.g., [290].

Green functions in layered poroelastic half-spaces can be calculated in the Laplace domain by using a propagator matrix method [218].

4.3 Boundary Integral Representations

In order to obtain Green's Third Identity, we also need fundamental solutions of the adjoint Eqs. (4.12). Strictly speaking, we need a system of functions satisfying

$$\begin{aligned}&\begin{pmatrix} -\frac{\lambda+\mu}{\mu}\nabla_y\left(\nabla_y\cdot v(y,\tau;x,t)\right) - \nabla_y^2 v(y,\tau;x,t) + \alpha\nabla_y\left(\partial_\tau q(y,\tau;x,t)\right) \\ -\partial_\tau\left(c_0\mu q(y,\tau;x,t)\right) - \alpha\nabla_y\cdot v(y,\tau;x,t) - \nabla_y^2 q(y,\tau;x,t) \end{pmatrix} \\ =&\boldsymbol{I}\delta(x-y)\delta(t-\tau)\ . \end{aligned} \tag{4.52}$$

Here, y is the spatial variable and τ is the time variable, whereas the singularity of the fundamental solutions is located at (x,t). With the observation

$$p^{\mathrm{Fi}}(x-y,t-\tau) = \partial_t u^{\mathrm{Si}}(x-y,t-\tau)\ , \tag{4.53}$$

which can be best seen by comparing (4.31) and (4.32), it can be shown that the fundamental solutions of (4.12) are

$$q^{\mathrm{Si}}(y,\tau;x,t) = p^{\mathrm{Si}}(x-y,t-\tau)\ , \tag{4.54a}$$

$$v^{\mathrm{Si}}(y,\tau;x,t) = p^{\mathrm{Fi}}(x-y,t-\tau)\ , \tag{4.54b}$$

$$q^{\mathrm{Fi}}(y,\tau;x,t) = u^{\mathrm{Si}}(x-y,t-\tau)\ , \tag{4.54c}$$

$$\boldsymbol{v}_{ik}^{\mathrm{Fi}}(y,\tau;x,t) = \boldsymbol{u}_{ki}^{\mathrm{Fi}}(x-y,t-\tau) \tag{4.54d}$$

which we summarize by

$$\boldsymbol{G}^*(y-x,\tau-t) = \begin{pmatrix} \boldsymbol{v}^{\mathrm{Fi}}(y-x,\tau-t) & v^{\mathrm{Si}}(y-x,\tau-t) \\ \left(q^{\mathrm{Fi}}(y-x,\tau-t)\right)^T & q^{\mathrm{Si}}(y-x,\tau-t) \end{pmatrix}\ . \tag{4.55}$$

Note that we have

$$\boldsymbol{G}^*(y-x,\tau-t) = (\boldsymbol{G}(x-y,t-\tau))^T \ . \tag{4.56}$$

We now have the ingredients to formulate boundary integral representations.

Theorem 4.11 (Boundary Integral Representations, Green's Third Identity) *Let $(u,p) \in C^1\left([0,t_{\text{end}}],\mathcal{C}^2(\overline{\Omega})\right) \times C^1\left([0,t_{\text{end}}],C^2(\overline{\Omega})\right)$ be a solution of (4.1) with vanishing right-hand sides. Then, for $(x,t) \in \Omega \times (0,t_{\text{end}})$, we formally have the boundary integral representations*

$$\begin{aligned}
u_i(x,t) = &\int_0^t \int_\Gamma \left(u_{ki}^{\text{Fi}}(x-y,t-\tau)\left[\boldsymbol{\sigma}_{kj,y}(u(y,\tau))n_j(y) - \alpha p(y,\tau)n_k(y)\right]\right. \\
&\qquad \left. + u_i^{\text{Si}}(x-y,t-\tau)\left[\left(\partial_{y_k} p(y,\tau)\right) n_k(y)\right]\right) \mathrm{d}S(y)\mathrm{d}\tau \\
&+ \int_0^t \left(\int_\Gamma \left[\boldsymbol{\sigma}_{kj,x}(\boldsymbol{u}_{\cdot i}^{\text{Fi}}(x-y,t-\tau))n_j(y)\right.\right. \\
&\qquad \left. - \alpha p_i^{\text{Fi}}(x-y,t-\tau)n_k(y)\right] u_k(y,\tau) \\
&\qquad \left. + \left[\left(\partial_{x_k} u_i^{\text{Si}}(x-y,t-\tau)\right) n_k(y)\right] p(y,\tau)\right) \mathrm{d}S(y)\mathrm{d}\tau \\
&+ \int_\Omega u_i^{\text{Si}}(y,0;x,t)\left(c_0\mu p(y,0) + \alpha\partial_{y_k} u_k(y,0)\right)\ \mathrm{d}V(y)\ , \qquad (4.57a)
\end{aligned}$$

$$\begin{aligned}
p(x,t) = &\int_0^t \int_\Gamma \left(p_k^{\text{Fi}}(x-y,t-\tau)\left[\boldsymbol{\sigma}_{kj,y}(u(y,\tau))n_j(y) - \alpha p(y,\tau)n_k(y)\right]\right. \\
&\qquad \left. + p^{\text{Si}}(x-y,t-\tau)\left[\left(\partial_{y_k} p(y,\tau)\right) n_k(y)\right]\right) \mathrm{d}S(y)\mathrm{d}\tau \\
&+ \int_0^t \int_\Gamma \left(\left[\partial_t\left(\boldsymbol{\sigma}_{kj,x}(u^{\text{Si}}(x-y,t-\tau))n_j(y)\right.\right.\right. \\
&\qquad \left.\left. \left. - \alpha p^{\text{Si}}(x-y,t-\tau)n_k(y)\right)\right] u_k(y,\tau)
\end{aligned}$$

$$+\left[\left(\partial_{x_k} p^{\mathrm{Si}}(x-y,t-\tau)\right) n_k(y)\right] p(y,\tau)\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau$$

$$+\int_{\Omega} p^{\mathrm{Si}}(y,0;x,t)\left(c_0\mu p(y,0)+\alpha\partial_{y_k}u_k(y,0)\right)\,\mathrm{d}V(y)\,. \tag{4.57b}$$

Here, $\boldsymbol{u}^{\mathrm{Fi}}_{\cdot i}$ is the i-th column of $\boldsymbol{u}^{\mathrm{Fi}}$.

Proof We start with Green's Second Identity (4.11) and choose v as well as q according to a column of $\boldsymbol{G}^*(y-x,\tau-t)$. On the left-hand side, the terms incorporating derivatives of u and p vanish as we assume that we have a solution of (4.1) with vanishing right-hand sides. The integral over Ω also vanishes as q vanishes for $\tau = t_{\mathrm{end}}$. The non-vanishing terms on the left-hand side reduce to a component of $u(x,t)$ or to $p(x,t)$, depending on which column of $\boldsymbol{G}^*(y-x,\tau-t)$ we choose, such that we obtain

$$\begin{aligned}
&u_i(x,t)\\
&=\int_0^t\int_{\Gamma}\Bigg(v^{\mathrm{Fi}}_{ki}(y,\tau;x,t)\left[\boldsymbol{\sigma}_{kj,y}(u(y,\tau))n_j(y)-\alpha p(y,\tau)n_k(y)\right]\\
&\qquad\qquad+q^{\mathrm{Fi}}_i(y,\tau;x,t)\left[\left(\partial_{y_k}p(y,\tau)\right)n_k(y)\right]\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&\quad-\int_0^t\int_{\Gamma}\Bigg(\Bigg[\Bigg(\frac{\lambda}{\mu}\partial_{y_l}v^{\mathrm{Fi}}_{li}(y,\tau;x,t)\delta_{jk}+\partial_{y_j}v^{\mathrm{Fi}}_{ki}(y,\tau;x,t)\\
&\qquad\qquad+\partial_{y_k}v^{\mathrm{Fi}}_{ji}(y,\tau;x,t)\Bigg)n_j(y)-\alpha\partial_\tau q^{\mathrm{Fi}}_i(y,\tau;x,t)n_k(y)\Bigg]u_k(y,\tau)\\
&\qquad\qquad+\left[\left(\partial_{y_k}q^{\mathrm{Fi}}_i(y,\tau;x,t)\right)n_k(y)\right]p(y,\tau)\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&\quad+\int_{\Omega}q^{\mathrm{Fi}}_i(y,0;x,t)\left(c_0\mu p(y,0)+\alpha\partial_{y_k}u_k(y,0)\right)\,\mathrm{d}V(y)\\
&=\int_0^t\int_{\Gamma}\Bigg(u^{\mathrm{Fi}}_{ik}(x-y,t-\tau)\left[\boldsymbol{\sigma}_{kj,y}(u(y,\tau))n_j(y)-\alpha p(y,\tau)n_k(y)\right]\\
&\qquad\qquad+u^{\mathrm{Si}}_i(x-y,t-\tau)\left[\left(\partial_{y_k}p(y,\tau)\right)n_k(y)\right]\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau
\end{aligned}$$

$$
\begin{aligned}
&- \int_0^t \int_\Gamma \Bigg(\Bigg[\Bigg(\Bigg(\frac{\lambda}{\mu} \partial_{y_l} u_{il}^{\mathrm{Fi}}(x-y,t-\tau)\delta_{jk} + \partial_{y_j} u_{ik}^{\mathrm{Fi}}(x-y,t-\tau) \\
&\qquad + \partial_{y_k} u_{ij}^{\mathrm{Fi}}(x-y,t-\tau) \Bigg) n_j(y) \\
&\qquad - \alpha \partial_\tau u_i^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \Bigg] u_k(y,\tau) \\
&\quad + \Big[\big(\partial_{y_k} u_i^{\mathrm{Si}}(x-y,t-\tau) \big) n_k(y) \Big] p(y,\tau) \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau \\
&+ \int_\Omega u_i^{\mathrm{Si}}(y,0;x,t) \left(c_0 \mu p(y,0) + \alpha \partial_{y_k} u_k(y,0) \right) \, \mathrm{d}V(y) \\
&= \int_0^t \int_\Gamma \Bigg(u_{ki}^{\mathrm{Fi}}(x-y,t-\tau) \Big[\boldsymbol{\sigma}_{kj,y}(u(y,\tau)) n_j(y) - \alpha p(y,\tau) n_k(y) \Big] \\
&\qquad + u_i^{\mathrm{Si}}(x-y,t-\tau) \Big[\big(\partial_{y_k} p(y,\tau) \big) n_k(y) \Big] \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau \\
&+ \int_0^t \int_\Gamma \Bigg(\Bigg[\Bigg(\Bigg(\frac{\lambda}{\mu} \partial_{x_l} u_{li}^{\mathrm{Fi}}(x-y,t-\tau)\delta_{jk} + \partial_{x_j} u_{ki}^{\mathrm{Fi}}(x-y,t-\tau) \\
&\qquad + \partial_{x_k} u_{ji}^{\mathrm{Fi}}(x-y,t-\tau) \Bigg) n_j(y) \\
&\qquad - \alpha \partial_t u_i^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \Bigg] u_k(y,\tau) \\
&\quad + \Big[\big(\partial_{x_k} u_i^{\mathrm{Si}}(x-y,t-\tau) \big) n_k(y) \Big] p(y,\tau) \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau \\
&+ \int_\Omega u_i^{\mathrm{Si}}(y,0;x,t) \left(c_0 \mu p(y,0) + \alpha \partial_{y_k} u_k(y,0) \right) \, \mathrm{d}V(y) \\
&= \int_0^t \int_\Gamma \Bigg(u_{ik}^{\mathrm{Fi}}(x-y,t-\tau) \Big[\boldsymbol{\sigma}_{kj,y}(u(y,\tau)) n_j(y) - \alpha p(y,\tau) n_k(y) \Big] \\
&\qquad + u_i^{\mathrm{Si}}(x-y,t-\tau) \Big[\big(\partial_{y_k} p(y,\tau) \big) n_k(y) \Big] \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau
\end{aligned}
$$

$$
\begin{aligned}
&+\int_0^t\int_\Gamma \Bigg(\Bigg[\Bigg(\frac{\lambda}{\mu}\partial_{x_l}u_{li}^{\mathrm{Fi}}(x-y,t-\tau)\delta_{jk}+\partial_{x_j}u_{ki}^{\mathrm{Fi}}(x-y,t-\tau)\\
&\qquad+\partial_{x_k}u_{ji}^{\mathrm{Fi}}(x-y,t-\tau)\Bigg)n_j(y)\\
&\qquad-\alpha p_i^{\mathrm{Fi}}(x-y,t-\tau)n_k(y)\Bigg]u_k(y,\tau)\\
&\quad+\Big[\big(\partial_{x_k}u_i^{\mathrm{Si}}(x-y,t-\tau)\big)\,n_k(y)\Big]p(y,\tau)\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&+\int_\Omega u_i^{\mathrm{Si}}(y,0;x,t)\left(c_0\mu p(y,0)+\alpha\partial_{y_k}u_k(y,0)\right)\,\mathrm{d}V(y)
\end{aligned}
\tag{4.58}
$$

or

$$
\begin{aligned}
&p(x,t)\\
&=\int_0^t\int_\Gamma\Bigg(v_k^{\mathrm{Si}}(y,\tau;x,t)\Big[\boldsymbol{\sigma}_{kj,y}(u(y,\tau))n_j(y)-\alpha p(y,\tau)n_k(y)\Big]\\
&\qquad+q^{\mathrm{Si}}(y,\tau;x,t)\Big[\big(\partial_{y_k}p(y,\tau)\big)\,n_k(y)\Big]\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&-\int_0^t\int_\Gamma\Bigg(\Bigg[\Bigg(\frac{\lambda}{\mu}\partial_{y_l}v_l^{\mathrm{Si}}(y,\tau;x,t)\delta_{jk}+\partial_{y_j}v_k^{\mathrm{Si}}(y,\tau;x,t)\\
&\qquad+\partial_{y_k}v_j^{\mathrm{Si}}(y,\tau;x,t)\Bigg)n_j(y)-\alpha\partial_\tau q^{\mathrm{Si}}(y,\tau;x,t)n_k(y)\Bigg]u_k(y,\tau)\\
&\quad+\Big[\big(\partial_{y_k}q^{\mathrm{Si}}(y,\tau;x,t)\big)\,n_k(y)\Big]p(y,\tau)\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&+\int_\Omega q^{\mathrm{Si}}(y,0;x,t)\left(c_0\mu p(y,0)+\alpha\partial_{y_k}u_k(y,0)\right)\,\mathrm{d}V(y)\\
&=\int_0^t\int_\Gamma\Bigg(p_k^{\mathrm{Fi}}(x-y,t-\tau)\Big[\boldsymbol{\sigma}_{kj,y}(u(y,\tau))n_j(y)-\alpha p(y,\tau)n_k(y)\Big]\\
&\qquad+p^{\mathrm{Si}}(x-y,t-\tau)\Big[\big(\partial_{y_k}p(y,\tau)\big)\,n_k(y)\Big]\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau
\end{aligned}
$$

$$
\begin{aligned}
&- \int_0^t \int_\Gamma \Bigg(\Bigg[\Bigg(\Bigg(\frac{\lambda}{\mu} \partial_{y_l} p_l^{\mathrm{Fi}}(x-y,t-\tau)\delta_{jk} + \partial_{y_j} p_k^{\mathrm{Fi}}(x-y,t-\tau) \\
&\qquad + \partial_{y_k} p_j^{\mathrm{Fi}}(x-y,t-\tau) \Bigg) n_j(y) \\
&\qquad - \alpha \partial_\tau p^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \Bigg] u_k(y,\tau) \\
&\quad + \Big[\big(\partial_{y_k} p^{\mathrm{Si}}(x-y,t-\tau) \big) n_k(y) \Big] p(y,\tau) \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau \\
&+ \int_\Omega p^{\mathrm{Si}}(y,0;x,t) \left(c_0 \mu p(y,0) + \alpha \partial_{y_k} u_k(y,0) \right) \, \mathrm{d}V(y) \\
&= \int_0^t \int_\Gamma \Bigg(p_k^{\mathrm{Fi}}(x-y,t-\tau) \Big[\boldsymbol{\sigma}_{kj,y}(u(y,\tau)) n_j(y) - \alpha p(y,\tau) n_k(y) \Big] \\
&\qquad + p^{\mathrm{Si}}(x-y,t-\tau) \Big[\big(\partial_{y_k} p(y,\tau) \big) n_k(y) \Big] \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau \\
&+ \int_0^t \int_\Gamma \Bigg(\Bigg[\Bigg(\Bigg(\frac{\lambda}{\mu} \partial_{x_l} p_l^{\mathrm{Fi}}(x-y,t-\tau)\delta_{jk} + \partial_{x_j} p_k^{\mathrm{Fi}}(x-y,t-\tau) \\
&\qquad + \partial_{x_k} p_j^{\mathrm{Fi}}(x-y,t-\tau) \Bigg) n_j(y) \\
&\qquad - \alpha \partial_t p^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \Bigg] u_k(y,\tau) \\
&\quad + \Big[\big(\partial_{x_k} p^{\mathrm{Si}}(x-y,t-\tau) \big) n_k(y) \Big] p(y,\tau) \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau \\
&+ \int_\Omega p^{\mathrm{Si}}(y,0;x,t) \left(c_0 \mu p(y,0) + \alpha \partial_{y_k} u_k(y,0) \right) \, \mathrm{d}V(y) \\
&= \int_0^t \int_\Gamma \Bigg(p_k^{\mathrm{Fi}}(x-y,t-\tau) \Big[\boldsymbol{\sigma}_{kj,y}(u(y,\tau)) n_j(y) - \alpha p(y,\tau) n_k(y) \Big] \\
&\qquad + p^{\mathrm{Si}}(x-y,t-\tau) \Big[\big(\partial_{y_k} p(y,\tau) \big) n_k(y) \Big] \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau
\end{aligned}
$$

$$
\begin{aligned}
&+\int_0^t\int_\Gamma \Bigg(\Bigg[\partial_t\Bigg(\frac{\lambda}{\mu}\partial_{x_l}u_l^{\mathrm{Si}}(x-y,t-\tau)\delta_{jk}+\partial_{x_j}u_k^{\mathrm{Si}}(x-y,t-\tau)\\
&\qquad+\partial_{x_k}u_j^{\mathrm{Si}}(x-y,t-\tau)\Bigg)n_j(y)\\
&\qquad-\alpha p^{\mathrm{Si}}(x-y,t-\tau)n_k(y)\Bigg)\Bigg]u_k(y,\tau)\\
&\quad+\Bigg[\left(\partial_{x_k}p^{\mathrm{Si}}(x-y,t-\tau)\right)n_k(y)\Bigg]p(y,\tau)\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&+\int_\Omega p^{\mathrm{Si}}(y,0;x,t)\left(c_0\mu p(y,0)+\alpha\partial_{y_k}u_k(y,0)\right)\,\mathrm{d}V(y)\,. \qquad (4.59)
\end{aligned}
$$

Here, we changed derivatives affecting fundamental solutions from derivatives with respect to y and τ to derivatives with respect to x and t and used (4.53) as well as the symmetry of $\boldsymbol{u}^{\mathrm{Fi}}$.

Remark 4.12 As mentioned above, (4.57a) and (4.57b) are only formal representations because they contain strong singularities and, thus, these integrals may not exist in general or may only be interpreted in the sense of principal values.

Moreover, if we allow for points x on the boundary of the domain Ω, the solid angle subtended at x by the boundary Γ of Ω has to be included as a factor on the left-hand sides of (4.57a) and (4.57b).

According to [58], we can deduce single- and double-layer potentials from (4.57a) and (4.57b) to be

$$
\begin{aligned}
u_i(x,t) &= \int_0^t\int_\Gamma\left(u_{ki}^{\mathrm{Fi}}(x-y,t-\tau)\phi_k(y,\tau)+u_i^{\mathrm{Si}}(x-y,t-\tau)\phi_4(y,\tau)\right)\mathrm{d}S(y)\mathrm{d}\tau\\
&\quad+\int_\Omega u_i^{\mathrm{Si}}(y,0;x,t)\phi^{(0)}(y)\,\mathrm{d}V(y)\,, \qquad (4.60a)
\end{aligned}
$$

$$
\begin{aligned}
p(x,t) &= \int_0^t\int_\Gamma\left(p_k^{\mathrm{Fi}}(x-y,t-\tau)\phi_k(y,\tau)+p^{\mathrm{Si}}(x-y,t-\tau)\phi_4(y,\tau)\right)\mathrm{d}S(y)\mathrm{d}\tau\\
&\quad+\int_\Omega p^{\mathrm{Si}}(y,0;x,t)\phi^{(0)}(y)\,\mathrm{d}V(y)\,, \qquad (4.60b)
\end{aligned}
$$

and

$$u_i(x,t) = \int_0^t \int_\Gamma \Bigg(\Big[\boldsymbol{\sigma}_{kj,x}(\boldsymbol{u}^{\mathrm{Fi}}_{\cdot i}(x-y,t-\tau)) n_j(y) - \alpha p_i^{\mathrm{Fi}}(x-y,t-\tau) n_k(y) \Big] \psi_k(y,\tau) + \Big[\partial_{x_k} u_i^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \Big] \psi_4(y,\tau) \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau + \int_\Omega u_i^{\mathrm{Si}}(y,0;x,t) \psi^{(0)}(y) \, \mathrm{d}V(y) \,, \tag{4.60c}$$

$$p(x,t) = \int_0^t \int_\Gamma \Bigg(\Big[\partial_t \big(\boldsymbol{\sigma}_{kj,x}(u^{\mathrm{Si}}(x-y,t-\tau)) n_j(y) - \alpha p^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \big) \Big] \psi_k(y,\tau) + \Big[\partial_{x_k} p^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \Big] \psi_4(y,\tau) \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau + \int_\Omega p^{\mathrm{Si}}(y,0;x,t) \psi^{(0)}(y) \, \mathrm{d}V(y) \,. \tag{4.60d}$$

Here, the values of the densities $\phi(y,\tau)$ and $\psi(y,\tau)$ are in $\mathbb{R}^4$, whereas values of $\phi^{(0)}(y)$ and $\psi^{(0)}(y)$ are in $\mathbb{R}$. Indices i,k take values in $\{1,2,3\}$. Note that the boundary integral operators which can be defined based on these potentials will not be self adjoint, because the differential operator of the underlying differential equations is already not self adjoint.

We have seen that there are relations of the QEP to the heat equation, the Cauchy-Navier equation, and the Stokes equations. These relations are reflected by the corresponding fundamental solutions. Thus, it is reasonable to assume that well known limit and jump relations are also valid for the layer potentials given above (see, e.g., [91, 194] and the references therein for the Laplace and the Cauchy-Navier equation, [57, 101] for the heat equation, as well as [191, 192] for the Stokes equations). A thorough investigation of the properties of these boundary layer potentials is out of the scope of this thesis. Such an investigation would include determining on which function spaces (e.g., on which Sobolev spaces or Hölder spaces) the associated operators are defined, proving continuity properties and determining whether limit and jump relations hold.

Remark 4.13 In engineering sciences, boundary integral equations are often deduced otherwise by starting with a so-called reciprocal theorem. For example, Cheng and Detournay (see [52]) derived boundary integral equations from a reciprocal theorem which holds in poroelasticity [136, 213, 227]

$$\boldsymbol{\sigma}_{ij}^{\mathrm{pe},(1)}\boldsymbol{\epsilon}_{ij}^{(2)} + p^{(1)}\zeta^{(2)} = \boldsymbol{\sigma}_{ij}^{\mathrm{pe},(2)}\boldsymbol{\epsilon}_{ij}^{(1)} + p^{(2)}\zeta^{(1)}$$

This is a generalization of Betti's reciprocal theorem that holds in classical linear elasticity [188]. Here, the superscripts (1) and (2) indicate "quantities under two independent stress and strain states at different spatial and time coordinates" [52]. The boundary integral equations given in [52] can be written (by using only the fundamental solutions we have introduced so far) as

$$\begin{aligned}
&u_i(x,t)\\
&= \int_0^t \int_\Gamma \Bigg(u_{ki}^{\mathrm{Fi}}(x-y,t-\tau)\Big[\boldsymbol{\sigma}_{kj,y}(u^{(1)}(y,\tau))n_j(y) - \alpha p^{(1)}(y,\tau)n_k(y)\\
&\qquad\qquad - \left(\boldsymbol{\sigma}_{kj,y}(u^{(2)}(y,\tau))n_j(y) - \alpha p^{(2)}(y,\tau)n_k(y)\right)\Big]\\
&\qquad + \left(\partial_{x_k} u_i^{\mathrm{Si}}(x-y,t-\tau)\right)\Big[\left(p^{(1)}(y,\tau) - p^{(2)}(y,\tau)\right)n_k(y)\Big]\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&\quad + \int_\Omega u_i^{\mathrm{Si}}(y,0;x,t)\left(c_0\mu p(y,0) + \alpha\partial_{y_k}u_k(y,0)\right)\,\mathrm{d}V(y) \qquad (4.61a)\\
&p(x,t)\\
&= \int_0^t \int_\Gamma \Bigg(p_k^{\mathrm{Fi}}(x-y,t-\tau)\Big[\boldsymbol{\sigma}_{kj,y}(u^{(1)}(y,\tau))n_j(y) - \alpha p^{(1)}(y,\tau)n_k(y)\\
&\qquad\qquad - \left(\boldsymbol{\sigma}_{kj,y}(u^{(2)}(y,\tau))n_j(y) - \alpha p^{(2)}(y,\tau)n_k(y)\right)\Big]\\
&\qquad + \left(\partial_{x_k} p^{\mathrm{Si}}(x-y,t-\tau)\right)\Big[\left(p^{(1)}(y,\tau) - p^{(2)}(y,\tau)\right)n_k(y)\Big]\Bigg)\,\mathrm{d}S(y)\mathrm{d}\tau\\
&\quad + \int_\Omega p^{\mathrm{Si}}(y,0;x,t)\left(c_0\mu p(y,0) + \alpha\partial_{y_k}u_k(y,0)\right)\,\mathrm{d}V(y). \qquad (4.61b)
\end{aligned}$$

for the stress discontinuity method and

$$
\begin{aligned}
&u_i(x,t)\\
&= \int_0^t \int_\Gamma \Bigg(\Big[\boldsymbol{\sigma}_{kj,x}(\boldsymbol{u}^{\mathrm{Fi}}_{\cdot i}(x-y,t-\tau)) n_j(y) - \alpha p_i^{\mathrm{Fi}}(x-y,t-\tau) n_k(y) \Big] \\
&\qquad \times \left(u_k^{(1)}(y,\tau) - u_k^{(2)}(y,\tau) \right) \\
&\qquad + \left[u_i^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \right] \left(\partial_{y_k} p^{(1)}(y,\tau) - \partial_{y_k} p^{(2)}(y,\tau) \right) \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau \\
&\quad + \int_\Omega u_i^{\mathrm{Si}}(y,0;x,t) \left(c_0 \mu p(y,0) + \alpha \partial_{y_k} u_k(y,0) \right) \, \mathrm{d}V(y)
\end{aligned}
\tag{4.62a}
$$

$$
\begin{aligned}
&p(x,t)\\
&= \int_0^t \int_\Gamma \Bigg(\Big[\partial_t \big(\boldsymbol{\sigma}_{kj,x}(u^{\mathrm{Si}}(x-y,t-\tau)) n_j(y) - \alpha p^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \big) \Big] \\
&\qquad \times \left(u_k^{(1)}(y,\tau) - u_k^{(2)}(y,\tau) \right) \\
&\qquad + \left[p^{\mathrm{Si}}(x-y,t-\tau) n_k(y) \right] \left(\partial_{y_k} p^{(1)}(y,\tau) - \partial_{y_k} p^{(2)}(y,\tau) \right) \Bigg) \, \mathrm{d}S(y) \mathrm{d}\tau \\
&\quad + \int_\Omega p^{\mathrm{Si}}(y,0;x,t) \left(c_0 \mu p(y,0) + \alpha \partial_{y_k} u_k(y,0) \right) \, \mathrm{d}V(y).
\end{aligned}
\tag{4.62b}
$$

for the displacement discontinuity method. As in the original article [52], we used index notation and summation convention. Here, we did not expand the integrals that include initial conditions over the whole space $\mathbb{R}^3$ as this is only necessary if a time stepping scheme is used (see [52] for further explanations).

The difference between those formulations and the single- and double-layer potentials which we derived is only in an additional integration by parts regarding the pressures. This accounts for the fact that, from a physical point of view, it is more consistent to describe either normal tensions and pressures or displacements and normal fluid flow. According to Darcy's Law (4.5), the latter is equivalent to the normal derivative of the pressure.

Although we now have boundary integral equations and, thus, reduced the dimension, we have not yet satisfied the aspiration to reduce the numerical costs. This is due to the fact that we would need to choose an ansatz, e.g., a single-layer with an unknown density, which would lead to some integral equations of which

we do not even know if they are Fredholm integral equations. A numerical solution scheme for a boundary layer approach is confronted with at least two difficulties:

(i) Even for rather simple boundaries, like a sphere, elaborate integration schemes are needed. Numerical integration over realistic boundaries requires even more sophisticated techniques, e.g., for meshing the boundary.
(ii) Moreover, we have to deal with a singular integrand. This increases the demands on the stability of the integration scheme even further.

Furthermore, to evaluate the solution at any time t, we would need to compute another integral over $\Gamma \times (0, t)$. This results in a numerical method which is approximately as expensive as a direct method in terms of memory and CPU time.

To overcome the drawbacks of a boundary integral equations formulation, we will introduce a less expensive numerical procedure which is related to layer potential approaches in the next chapter.

Chapter 5
Methods of Fundamental Solutions in Poroelasticity

The method of fundamental solutions, which we will abbreviate as MFS, is also known in mathematical terminology as the superposition method, the desingularized method, the fundamental collocation method, and the charge simulation method (see [120] and the references therein). At least in the field of physics it is best known as the method of image charges in electrostatics (see, e.g., [137] and the references therein).

The aim of this chapter is to develop a method of fundamental solutions in quasistatic poroelasticity beginning with a single-layer potential approach. We will summarize some results on the density of fundamental solutions as ansatz functions in case of elliptic equations and on convergence properties in some special cases. Furthermore, we will show density of the fundamental solutions of poroelasticity for a problem with vanishing initial values. Incorporating a non-vanishing initial condition leads to an alternative approach on how to use fundamental solutions in the context of quasistatic poroelasticity. Numerical simulations will be presented in the next chapter.

Some of the contents of this chapter, especially of the first section, were already published before in [14] and some of the basic ideas in [17].

5.1 The Method of Fundamental Solutions: An Overview

One of the most popular numerical approaches to find a solution of a partial differential equation is the finite element method. It is based on the Galerkin method which is a generalization of the Ritz method, first formulated by Ritz in 1909 [236] as an application of the Riesz Representation Theorem 2.58 and the Lax-Milgram Theorem 2.59.

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4_5

As we have seen before, a partial differential equation in its weak form can be formulated using a bilinear form $a(\cdot, \cdot)$ defined on a Hilbert space V. The right hand side f of the equation is regarded as an element of the dual space V'. A weak solution of the differential equation is then defined as a solution of

$$a(u, v) = f(v) \quad \forall v \in V \,. \tag{5.1}$$

Alternatively, the solution can be found by finding the minimizer u to the minimization problem

$$F(v) = \frac{1}{2} a(v, v) - f(v) \to \min \,. \tag{5.2}$$

Let $(\phi_n)_{n\in\mathbb{N}}$ be a basis of V. The solution u for which we are looking can be expressed as a series with respect to these basis functions. For a numerical method, it is necessary to approximate u by a finite sum of these basis functions, i.e., by $u_N = \sum_{n=1}^{N} a_n \phi_n$. In the context of Galerkin methods, it is convenient to choose such kind of basis functions that given (Dirichlet) boundary conditions are automatically satisfied by the ansatz. To determine the coefficients a_n, the approximate solution is plugged in the bilinear form as u and the basis functions used in the approximation are plugged in as v. This results in a regular system of linear equations. Therefore, u_n is the solution of (5.1) not in the whole of V but in the subspace spanned by $(\phi_n)_{n=0}^{N}$.

In 1926, Trefftz [274] suggested an alternative to the Ritz method. Instead of using basis functions that satisfy the boundary conditions and compute the coefficients as mentioned above, he suggested to use basis functions that each satisfy (5.1) with $f = 0$ and to determine the coefficients via the condition

$$\left\| \sum_{n=1}^{N} a_n \phi_n - g \right\| \to \min \tag{5.3}$$

with $\|\cdot\|$ being a norm that has to be specified and g the given boundary data. Here, we assumed Dirichlet boundary data given on the whole boundary Γ. For a more general approach, see, e.g., [129]. An overview on the Trefftz method can be found, e.g., in [158]. There, it is also shown how the method of fundamental solutions (MFS) can be interpreted as a modified indirect Trefftz method (see also [50]).

Instead of just formulating a Trefftz method using fundamental solutions, we motivate this ansatz by showing how the MFS can be seen as a regularized single-layer approach. For this purpose, we consider the Laplace equation with a Dirichlet boundary condition on the whole boundary Γ, following [116] and for the sake of simplicity of presentation of the basic ideas (see also [80] as well as [91] and the references therein for further details on the properties of a system of fundamental solutions to approximate harmonic functions). The basic idea is the same as used by Runge [239] or Walsh [284] to prove that holomorphic functions in an open domain

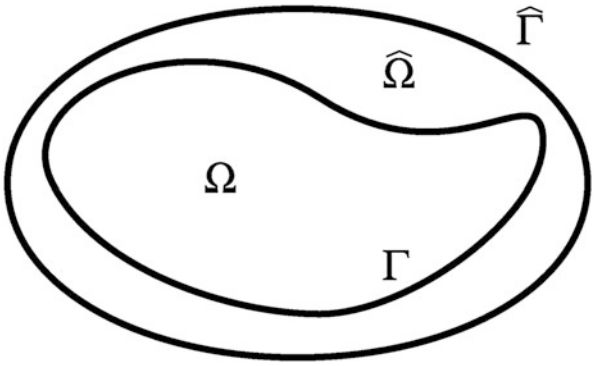

Fig. 5.1 Domain Ω with boundary Γ and $\hat{\Omega} \supset \Omega$ with boundary $\hat{\Gamma}$

U of the complex plane can be approximated uniformly in compact subsets of U by rational functions: transform a singular integral into a regular integral by moving the singularity away from the domain of interest. In order to do this, we introduce a larger domain $\hat{\Omega} \supset \Omega$ with boundary $\hat{\Gamma}$ (Fig. 5.1) and assume that the solution u can be written as a single-layer potential with unknown density and integration over $\hat{\Gamma}$ such that

$$u(x) = \int_{\hat{\Gamma}} \phi(y) G^{\text{Harm}}(x - y) \, \mathrm{d}S(y) \, . \tag{5.4}$$

On the boundary Γ of Ω, the boundary condition has to be satisfied, thus,

$$g(x) = u(x)|_{\Gamma} = \int_{\hat{\Gamma}} \phi(y) G^{\text{Harm}}(x - y) \, \mathrm{d}S(y) \quad \text{for all } x \in \Gamma \, . \tag{5.5}$$

Let $(\psi_j)_{j \in \mathbb{N}}$ be a system of functions which can be defined at least on a domain slightly larger than $\hat{\Omega}$ and including $\hat{\Omega}$ such that $(\psi_j|_{\hat{\Gamma}})_{j \in \mathbb{N}}$ is a complete system of functions on $\hat{\Gamma}$. Then $\phi(y)$ can be approximated by a finite sum $\phi_J(y) = \sum_{j=1}^{J} c_j \psi_j(y)$, $J \in \mathbb{N}$. Moreover, we only demand that (5.5) has to be satisfied in some discrete collocation points $(x_{(i)})_{i=1}^{I}$, $I \in \mathbb{N}$, along the boundary Γ. We put the index here in parentheses to avoid confusion with the i-th component of the vector x. With these assumptions, we obtain

$$g(x_{(i)}) = \sum_{j=1}^{J} c_j \int_{\hat{\Gamma}} \psi_j(y) G^{\text{Harm}}(x_{(i)} - y) \mathrm{d}S(y) \quad \text{for all } 1 \le i \le I \, . \tag{5.6}$$

Finally, we take into consideration that we are talking about a numerical method. Thus, we will use a cubature rule with weights w_m to compute the integrals above

which yields

$$g(x_{(i)}) = \sum_{j=1}^{J} c_j \sum_{m=1}^{M} w_m \psi_j(y_{(m)}) G^{\mathrm{Harm}}(x_{(i)}, y_{(m)})$$
$$= \sum_{m=1}^{M} \underbrace{w_m \sum_{j=0}^{J} c_j \psi_j(y_{(m)})}_{=a_m} G^{\mathrm{Harm}}(x_{(i)}, y_{(m)}) = \sum_{m=1}^{M} a_m G^{\mathrm{Harm}}(x_{(i)}, y_{(m)}) \tag{5.7}$$

with $M \in \mathbb{N}$. This result leads to the conclusion that we can alternatively choose an ansatz for the solution u by

$$u(x) \approx u_M(x) = \sum_{m=1}^{M} a_m G^{\mathrm{Harm}}(x, y_{(m)}) \,, \tag{5.8}$$

i.e., the solution is approximated as a finite sum of fundamental solutions of the differential equation under consideration with poles $(y_{(m)})_{m=1}^{M}$ outside the domain on which we are interested in this solution. For the sake of simplicity in notation, we further on drop the distinction between u and an approximation to u.

On Ω, every function $G^{\mathrm{Harm}}(\,\cdot\,, y_{(m)})$, $m \in \mathbb{N}$, $1 \leq m \leq M$, is analytic and satisfies $\nabla_x^2 G^{\mathrm{Harm}}(x, y_{(m)}) = 0$. Hence, Ansatz (5.8) follows the ideas of Trefftz as mentioned above.

Using an ansatz like (5.8) right from the beginning, it is clear that we do not have to restrict ourselves to pure Dirichlet boundary value problems but can also address other, e.g., mixed boundary value problems, or, in case of time-dependent functions, initial boundary value problems. Furthermore, instead of using collocation to approximate the boundary, a least-squares approximation may be used. Moreover, we do not need to perform a numerical integration over the boundary explicitly. Thus, the MFS can be implemented as an integration-free, meshless method, granting several benefits [262]:

- As the MFS is a meshless method, we are totally free in choosing the points at which the corresponding singularities are located, in space as well as in time. We will call these points "source points".
- We are also free in the choice of the collocation points.
- We do not need an elaborate discretization (meshing) of the boundary or the domain. Thus, it can be applied to domains with irregular boundaries.
- By working only on the boundary, the dimension of the problem is reduced by one.
- There is no evaluation of (boundary) integrals incorporated, which usually requires elaborate integration techniques.
- The MFS requires little data preparation.

- The approximate solution and its derivatives can be evaluated directly at any point in the interior of the domain, whereas by using other boundary integral methods, such as the boundary element method, another integration is necessary to do this.
- The MFS can be easily implemented. Because of the above mentioned benefits, memory demands and CPU time required to solve a given problem are significantly reduced.

Nevertheless, there are some drawbacks of the MFS:

- The choice of collocation and source points can be crucial for a good performance, but there is, in general, very little advice on how to choose them or whether to choose the same number or different numbers of source and collocation points. For discussions on these problems, see [8, 24, 131, 155, 246, 266, 267]. In a real world problem, collocation points cannot be chosen freely, as data is only available at certain predefined points.
- The condition of the matrices can be very poor [262].
- A general convergence result is still unknown and results on the density of systems of fundamental solutions depend heavily on the differential equation under consideration as well as on how the set of source points is chosen.

The following paragraphs give an (incomplete) overview on the application of the MFS and mention some major theoretical results.

As mentioned above, some of the ideas of the MFS can be dated back to Runge in 1885 and Trefftz in 1926. As a method of its own, the MFS was first formulated by Aleksidze and Kupradze in the early to mid-1960s [4, 163–167]. In these articles, the MFS was already proposed as a solution method for problems of potential theory, elastostatics, and heat transport. First convergence results for the Laplace equation can be found in [167]. Even before that, in 1962, Browder [41] considered linear combinations of fundamental solutions for elliptic operators L with C^∞-coefficients. He proved that if one takes the fundamental solutions for a set of source points which is dense in an arbitrary open set outside the closure of an open domain Ω without holes and satisfying the cone condition, the set of their linear combinations is dense in the space

$$X = \{u \in C^m(\Omega) : \ Lu = 0 \text{ in } \Omega\} \cap C(\overline{\Omega})$$

with respect to the C^0-norm. This result was extended by Weinstock [287] to less regular solutions and domains which have the segment property.

Mathon and Johnston proposed and examined an MFS in which the position of the source points is variable for elliptic problems [189]. This leads to a non-linear optimization problem. Therefore, in most applications, the position of the source points is fixed.

Freeden [79] considered the external gravitational field and proved density of the fundamental solutions to the Laplace equations for solutions of an external Dirichlet problem. Müller and Kersten [206] showed density of the fundamental solutions of the Helmholtz equation in $\mathbb{R}^3 \setminus \overline{\Omega}$ up to eigensolutions in $L^2(\partial\Omega)$. Freeden and Kersten [89], who investigated the oblique derivative problem in potential theory, proved a density result for Hölder-spaces on the boundary and Freeden [80], who considered least-squares approximation by (multi-)poles, showed density in $L^2(\partial\Omega)$ for the Laplace equation. In all of these articles, the exterior of a bounded open domain Ω is considered and the source points are assumed to be located on the boundary of a smaller domain contained in Ω.

Bogomolny [40] showed density for the MFS in case of the Laplace, biharmonic, and modified Helmholtz equation in two and three dimensions as well as convergence in case of the Laplace equation in two dimensions for some special choice of source points on a circle in case of domains whose boundary lies between two circles with respect to the supremum norm. Further on, Katsurada and Okamoto improved Bogomolny's convergence results to more general domains [151–154], while Kitagawa considered the stability of the MFS for the Laplace equation on disks [159, 160]. Smyrlis and Karageorghis analyzed convergence and stability of the MFS for the Laplace equation on disks by considering the properties of the arising matrices [265] and, together with Tsangaris, on annular domains [280]. Whereas all of the above convergence results consider collocation methods, Smirlys proved convergence of the MFS for the Laplace equation on disks in case of a least-squares approach by using weighted norms on certain Sobolev spaces [261]. In case of Poisson's equation, Li used particular solutions based on radial basis functions and showed convergence of these together with an MFS-approach on disks in two dimensions [172] and on the unit ball in three dimensions [173]. Li extended his results to the modified Helmholtz equation on the unit ball with non-vanishing right-hand side [174]. For Helmholtz problems, Barnett and Betcke established convergence and stability results on analytic domains with a special choice of source points [24]. Grothaus and Raskop [121, 232] considered limit and jump relations of potential theory in Sobolev spaces and also showed density of the fundamental solutions to the Laplace equation in such a setting for oblique derivative problems.

Most of the results above consider elliptic equations, although Weinstock [287] briefly discussed the case of non-elliptic operators. But there are also density results for the heat equation in three dimensions by Kupradze [163] as well as in one and two dimensions by Johansson, Lesnic, and Reeve [140, 143]. Here, density of linear combinations of fundamental solutions is shown with respect to $L^2(\partial\Omega \times (0, t_{\text{end}}))$ and $L^2(\Omega \times \{0\})$, where $\Omega \in \mathbb{R}^n$, $n \in \{1, 2, 3\}$.

There are also density results for systems of differential equations. For example, Alves and Silvestre showed density of the fundamental solutions of the system of Stokes equations, so-called stokeslets, for interior and exterior problems in two and three dimensions [10]. Density results for the Cauchy-Navier equation and the system of equations of static thermoelasticity are given by Smyrlis [264].

As can already be seen by the articles cited above, the MFS has been applied to a variety of fields since its introduction. We can only give a few examples by mentioning applications to the fields of potential theory, potential flow, and Stokes flow [47, 145, 278, 279, 299, 302], the biharmonic equation [147], the Helmholtz equation [24], the modified Helmholtz equation [187], elastostatics [76, 148, 150, 196, 233], Signorini problems [226], fracture mechanics [124, 149], the wave equation and acoustics [12, 122, 162], heat conduction [43, 140, 141, 143], diffusion[1] [46, 133, 296, 297, 300], Stefan problems [44, 142], Brinkman flows [275], oscillatory and porous buoyant flow [277], diffusion-reaction equations [22], calculation of eigenfrequencies and eigenmodes [9], radiation and scattering problems [75], acoustic wave scattering on poroelastic scatterers [211], microstrip antenna analysis [252], or to two-dimensional unsteady Burger's equations[2][298].

Recently, Wen and Liu [288] used an MFS in poroelasticity. For this purpose, they used a formulation based on displacement vector for the solid and velocity of the fluid relative to the solid rather than our formulation which uses a displacement vector for the solid and the pressure of the fluid. Moreover, they considered the dynamic equations and applied a Laplace transform to get a set of time-independent equations.

For further examples, see the review articles by Fairweather and Karageorghis [74], Fairweather, Karageorghis and Martin [75], as well as Golberg and Chen [116] and the references therein. A more recent overview is given in [45].

Although the titles of [195, 276] claim to deal with thermoelasticity, neither of these articles consider the full coupled equations of thermoelasticity, but treat the temperature gradient which arises in the equation governing the displacement as a body-force term, therefore, using a dual-reciprocity method to address this non-vanishing right-hand side. To the best of our knowledge, there are no published results considering the coupled system of equations governing thermoelasticity with heat conduction described by the heat equation.

5.2 Density Results for the MFS in the Case of Poroelasticity

The same chain of reasoning as used above for the Laplace equation can be used in quasistatic poroelasticity to obtain an MFS from the single-layer potential given in (4.60a) and (4.60b), respectively. For a better understanding, we give a form of the single-layer in which the fundamental solutions are split into one part containing Dirac's delta distribution with respect to time and another one without such delta distributions. Interpreting integration over said delta distributions as evaluation at a

[1]The diffusion equation is basically the same as the heat equation. Both can be converted into a Helmholtz equation by applying Fourier transformation with respect to time.

[2]Burger's equation is non-linear when using spatial coordinates, but can be transferred into an uncoupled system of scalar diffusion equations by using material coordinates (cf. Definition 2.50).

certain point in time, the single-layer reads

$$u_i(x,t) = \int_0^t \int_\Gamma u_{ki}^{\mathrm{fi}}(x-y,t-\tau)\phi_k(y,\tau) + u_i^{\mathrm{Si}}(x-y,t-\tau)\phi_4(y,\tau)\,\mathrm{d}S(y)\mathrm{d}\tau + \int_\Gamma u_{ki}^{\mathrm{CN}}(x-y)\phi_k(y,t)\,\mathrm{d}S(y) + \int_\Omega u_i^{\mathrm{Si}}(y,0;x,t)\phi^{(0)}(y)\,\mathrm{d}V(y)\,, \tag{5.9a}$$

$$p(x,t) = \int_0^t \int_\Gamma p_k^{\mathrm{fi}}(x-y,t-\tau)\phi_k(y,\tau) + p^{\mathrm{Si}}(x-y,t-\tau)\phi_4(y,\tau)\,\mathrm{d}S(y)\mathrm{d}\tau + \int_\Gamma p_k^{\mathrm{St}}(x-y)\phi_k(y,t)\,\mathrm{d}S(y) + \int_\Omega p^{\mathrm{Si}}(y,0;x,t)\phi^{(0)}(y)\,\mathrm{d}V(y)\,. \tag{5.9b}$$

such that $0 < t < t_{\mathrm{end}}$ with $u^{\mathrm{fi}}(x-y,t-\tau)$, $p^{\mathrm{fi}}(x-y,t-\tau)$ given by (4.39) and $p^{\mathrm{St}}(x-y)$ as defined in (4.34). From these, we obtain

$$\begin{aligned} u_i(x,t) &= \sum_{m=1}^{M}\left(\sum_{n=1}^{N}\left(\sum_{k=1}^{3} a_{nm}^{(k)}\boldsymbol{u}_{ki}^{\mathrm{fi}}(x-y_{(m)},t-\tau_n) + b_{nm}u_i^{\mathrm{Si}}(x-y_{(m)},t-\tau_n)\right)\right. \\ &\qquad \left. + \sum_{k=1}^{3} a_m^{\mathrm{CN},(k)}(t)\boldsymbol{u}_{ki}^{\mathrm{CN}}(x-y_{(m)})\right) \\ &\quad + \sum_{l=1}^{L} b_l^{(0)}u_i^{\mathrm{Si}}(x-y_l,t-\tau_{-1})\,, \end{aligned} \tag{5.10a}$$

$$\begin{aligned} p(x,t) &= \sum_{m=1}^{M}\left(\sum_{n=1}^{N}\left(\sum_{k=1}^{3} a_{nm}^{(k)}p_k^{\mathrm{fi}}(x-y_{(m)},t-\tau_n) + b_{nm}p^{\mathrm{Si}}(x-y_{(m)},t-\tau_n)\right)\right. \\ &\qquad \left. + \sum_{k=1}^{3} a_m^{\mathrm{CN},(k)}(t)p_k^{\mathrm{St}}(x-y_{(m)})\right) \\ &\quad + \sum_{l=1}^{L} b_l^{(0)}p^{\mathrm{Si}}(x-y_l,t-\tau_{-1})\,, \end{aligned} \tag{5.10b}$$

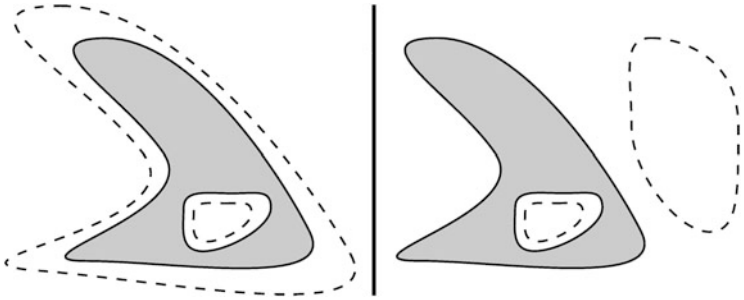

Fig. 5.2 Examples for an embracing pseudo-boundary (*dashed black lines*) to the domain Ω (*gray*) with boundary Γ (*solid black lines*). On the right-hand side, the unbounded component of $\mathbb{R}^n \setminus \overline{\Omega}$ contains a bounded component of $\mathbb{R}^n \setminus \hat{\Omega}$, which is not the case on the left-hand side (cf. [262, 264])

with $N, M, L \in \mathbb{N}$. Incorporating τ_{-1} in the sums approximating the integrals that incorporate initial values is necessary in order to allow for a collocation method to be used to approximate these initial values. For further details, see Sect. 5.2.2.

Although (5.10) is more suitable for implementation and (5.9) appears less artificial than a formulation involving delta distributions, the assumption of an ansatz consisting of a linear combination of the fundamental solutions given in (4.36) which includes delta distributions is more suitable to prove some theoretical results.

The aim of this section is to prove a density theorem for certain sets of fundamental solutions of the QEP in certain function spaces. For this purpose, it is useful to define the concept of an embracing pseudo-boundary [262].

Definition 5.1 Let Ω be an open connected subset of $\mathbb{R}^n$. We say the open connected subset $\hat{\Omega} \subset \mathbb{R}^n$ embraces Ω if $\overline{\Omega} \subset \hat{\Omega}$ and for every connected component U of $\mathbb{R}^n \setminus \overline{\Omega}$, there is an open connected component $\tilde{U}$ of $\mathbb{R}^n \setminus \overline{\hat{\Omega}}$ such that $\overline{\tilde{U}} \subset U$.

The boundary $\hat{\Gamma} = \partial\hat{\Omega}$ is called an embracing pseudo-boundary.

Figure 5.2 illustrates how embracing pseudo-boundaries may look like.

With Definition 5.1, we can now define the set of fundamental solutions under consideration.

5.2.1 *Density Results in the Case of Vanishing Initial Conditions*

For the sake of simplicity, we assume for now that the equations are equipped with homogeneous initial conditions. Therefore, the last integrals in (5.9), which incorporate initial conditions according to (4.57a) and (4.57b), should vanish for all

times $t \geq 0$. Consequently, the sums with index $l \in \mathbb{N}, 1 \leq l \leq L$, in (5.10) also vanish for all times $t \geq 0$.

Let $(y, \tau) \in \hat{\Gamma} \times (0, t_{\text{end}})$. We consider the set of all finite linear combinations of the form

$$\begin{aligned}\begin{pmatrix} u(x,t) \\ p(x,t) \end{pmatrix} &= \sum_{(y,\tau)} \boldsymbol{G}(x-y, t-\tau) \begin{pmatrix} a^{(1)}(y,\tau) \\ a^{(2)}(y,\tau) \\ a^{(3)}(y,\tau) \\ b(y,\tau) \end{pmatrix} \\ &= \sum_{(y,\tau)} \left(\boldsymbol{G}^*(y-x, \tau-t)\right)^T \begin{pmatrix} a^{(1)}(y,\tau) \\ a^{(2)}(y,\tau) \\ a^{(3)}(y,\tau) \\ b(y,\tau) \end{pmatrix} \end{aligned} \tag{5.11}$$

according to (5.10a) and (5.10b). Under the assumption $\operatorname{dist}(\overline{\Omega}, \hat{\Gamma}) > 0$, we obtain from (4.36) and Remark 4.9 that $\boldsymbol{G}_{ik}(x-y, t-\tau)|_{(x,t)\in\overline{\Omega}\times[0,t_{\text{end}}]} \in H^{-1}\left([0, t_{\text{end}}], C^l(\overline{\Omega})\right)$, $l \in \mathbb{N}$. Consider functionals $\nu_i \in \left(H^{-1}\left([0, t_{\text{end}}], C^l(\overline{\Omega})\right)\right)'$, $i \in \{1, 2, 3, 4\}$, such that

$$\nu(u, p) = \nu_1(u_1) + \nu_2(u_2) + \nu_3(u_3) + \nu_4(p) \, . \tag{5.12}$$

The action of ν on $\boldsymbol{G}$ is given by

$$\left(\nu(\tau_y \, \tau_\tau \, \boldsymbol{G})\right)_i = \nu_k(\tau_y \, \tau_\tau \, \boldsymbol{G}_{ki}) = \left(\nu_k * \boldsymbol{G}^*_{ik}\right)(y, \tau) = \left(\boldsymbol{G}^* * \nu\right)_i (y, \tau) \, . \tag{5.13}$$

Please remember that $(\tau_y f)(x) = f(x-y)$ and that we use the summation convention. This formula also explains how we understand the convolution of a matrix-valued and a vector-valued distribution, in the sense of the definition given by Hörmander [132, Definition 4.2.2], which is valid if at least one of the distributions has compact support.

The introduction of the above notations and definitions is motivated by three recent papers by Smyrlis [262–264]. In these papers, density results for elliptic equations on domains which have the segment property (see Definition 2.14) but may possess holes are proven. Here, a domain is regarded as possessing no holes if its complement is connected. We take a closer look at Smyrlis' papers now in order to show similar results for the QEP (4.1). The results in these three articles are based on the following lemma [262–264].

Lemma 5.2 *Let L be an elliptic operator with constant coefficients in $\mathbb{R}^n$ and $e = e(x)$ be a fundamental solution of L. Also, let Ω be an open bounded subset of $\mathbb{R}^n$ which has the segment property and ν a continuous linear functional on $C^l(\Omega)$, $l \in \mathbb{N}$. If $\vartheta = e * \nu$ is the convolution of the distributions e and ν and if* $\operatorname{supp}(\vartheta) \subset \overline{\Omega}$*, then there exists a sequence $\{\vartheta^k\}_{k\in\mathbb{N}} \subset (C_0^\infty(\Omega))'$ with $\operatorname{supp}(\vartheta^k) \subset \overline{\Omega}$ and $\{L\vartheta^k\}_{k\in\mathbb{N}} \subset (C^l(\Omega))'$, such that $\{L\vartheta^k\}_{k\in\mathbb{N}}$ converges to ν in the weak-$*$ sense of*

$(C^l(\Omega))'$, i.e., for every $u \in C^l(\Omega)$,

$$\lim_{k\to\infty} L\vartheta^k(u) = v(u) \ . \tag{5.14}$$

Remark 5.3 A similar result holds if v is a continuous linear functional on the space of functions with uniformly Hölder-continuous derivatives of order $k \in \mathbb{N}$, see [263].

Based on this lemma, Smyrlis proved a theorem for the MFS for elliptic differential equations similar to the one by Browder [41] which we mentioned in the previous section [263]. Unfortunately, this result is not exactly what we are looking for, because in this theorem a set of source points which is dense in an arbitrary, but fixed, open set outside the closure of an open domain is needed. This would imply that, in a numerical method, we take an open set and a mesh within this set. As the mesh becomes finer, the difference between the analytical and the numerical approximate solution decreases. Although such a method could be implemented, it would corrupt our efforts to reduce the costs of the numerical approach by reducing the dimension. Therefore, it would be preferable that, instead of a dense point set in an open set, a dense point set on the boundary of an open set would be sufficient. In general, this is not the case. However, in some cases, such as for the Laplace equation [40, 262], biharmonic equation [262], Cauchy-Navier equation [264], or the heat equation [143, 163], such a result – or a slightly modified one – can be proven.

The first thing we prove is that a proposition similar to Lemma 5.2 also holds in case of the QEP. The proof follows the ones given in [262–264]. We split it up into several steps and start by proving an estimate for the partial derivatives of the fundamental solutions of (4.1).

Lemma 5.4 *The fundamental solutions p^{Si}, u^{Si}, p^{fi}, and $\boldsymbol{u}^{\mathrm{fi}}$, and their derivatives satisfy for $t > \tau$*

$$\left|D_x^a \partial_t^b p^{\mathrm{Si}}(x-y,t-\tau)\right| \leq C(t-\tau)^{-\frac{1}{2}(3+|a|+2b)} \exp\left(-\tilde{C}\frac{\|x-y\|^2}{t-\tau}\right) , \tag{5.15}$$

$$\begin{aligned}\left|D_x^a \partial_t^b u_k^{\mathrm{Si}}(x-y,t-\tau)\right| \leq & C(t-\tau)^{-\frac{1}{2}(3+|a|+2b)} \exp\left(-\tilde{C}\frac{\|x-y\|^2}{t-\tau}\right) \\ & + \hat{C}\,\|x-y\|^{-2-|a|}\,\partial_t^b H(t-\tau) \ ,\end{aligned} \tag{5.16}$$

$$\left|D_x^a \partial_t^b p_k^{\mathrm{fi}}(x-y,t-\tau)\right| \leq C(t-\tau)^{-\frac{1}{2}(5+|a|+2b)} \exp\left(-\tilde{C}\frac{\|x-y\|^2}{t-\tau}\right) , \tag{5.17}$$

$$\begin{aligned}\left|D_x^a \partial_t^b u_{ki}^{\mathrm{fi}}(x-y,t-\tau)\right| \leq & C(t-\tau)^{-\frac{1}{2}(4+|a|+2b)} \exp\left(-\tilde{C}\frac{\|x-y\|^2}{t-\tau}\right) \\ & + \hat{C}\,\|x-y\|^{-3-|a|}\,\partial_t^b H(t-\tau) \ ,\end{aligned} \tag{5.18}$$

with $D_x^a = \partial_{x_1}^{a_1}\partial_{x_2}^{a_2}\partial_{x_3}^{a_3}$, $a \in \mathbb{N}^3$, $|a| = a_1 + a_2 + a_3$, $k, i \in \{1, 2, 3\}$, *and* $C, \tilde{C}, \hat{C} \in \mathbb{R}$ *constants. Moreover, we have for* $\boldsymbol{u}^{\mathrm{CN}}(x-y)$ *and* $p^{\mathrm{St}}(x-y)$

$$\left|D_x^a u_{ki}^{\mathrm{CN}}(x-y)\right| \leq C\,\|x-y\|^{-1-|a|}\,, \tag{5.19}$$

$$\left|D_x^a p_k^{\mathrm{St}}(x-y)\right| \leq C\,\|x-y\|^{-2-|a|}\,. \tag{5.20}$$

Proof For the heat kernel, the same estimate as stated above for p^{Si} is well known, see, e.g., [101, Chapter 9, Theorem 8]. With this and the obvious estimate for derivatives of G^{Harm}, the estimate for u_k^{Si} is obtained from (4.31) where the Heaviside function H is formally defined as the corresponding integral of Dirac's delta distribution δ. The estimate for p_k^{fi} follows from (4.53). The estimate for $\boldsymbol{u}^{\mathrm{fi}}$ results from the relation $\boldsymbol{u}_{ki}^{\mathrm{fi}}(x,t) = C_1\partial_{x_k} u_i^{\mathrm{Si}}(x,t)$. Estimates for p^{St} and $\boldsymbol{u}^{\mathrm{CN}}$ are obtained from obvious estimates for G^{Harm}.

This lemma directly yields a corollary about the integrability of the fundamental solutions of (4.1) in equivalence to the integrability of partial derivatives of the fundamental solutions of elliptic differential equations.

Corollary 5.5 *The fundamental solutions* p^{Si}, u^{Si}, p^{Fi}, *and* $\boldsymbol{u}^{\mathrm{Fi}}$, *and their first partial derivatives are elements of* $\overline{L^1_{\mathrm{loc}}((\mathbb{R}^3 \setminus B_\varepsilon(y)) \times [\tau, \infty))}$, $\varepsilon > 0$, $0 \leq \tau \leq t$, *whereas the closure is taken with respect to the weak topology on* L^1_{loc}, *i.e., in the sense of distributions.*

Proof This results directly from the estimates given in Lemma 5.4.

Next, we merge two other lemmata by Smyrlis [262].

Lemma 5.6 *Let* $\nu \in \left(\left(H^{-1}\left([0, t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4$ *and* $\chi \in C_0^\infty\left(\overline{\Omega} \times [0, t_{\mathrm{end}}]\right)$. *Let* $\boldsymbol{G}^*$ *be the matrix of fundamental solutions of the adjoint equations of poroelasticity with singularity in* $\hat{\Gamma} \times (0, t_{\mathrm{end}})$ *as defined above. If the distributions* ν_i *and* $\left(\boldsymbol{G}^* * \nu\right)_i$, $i \in \{1, 2, 3, 4\}$, *have compact support and the support of* $\left(\boldsymbol{G}^* * \nu\right)_i$ *is a subset of* $U_1 \times [0, t_{\mathrm{end}}]$, *with* U_1 *being a neighborhood of* Ω, $\Omega \subset U_1 \subset \hat{\Omega}$, $\mathrm{dist}(U_1, \hat{\Gamma}) > 0$, *then* $(L^{\mathrm{pe}})^*\left(\chi\left(\boldsymbol{G}^* * \nu\right)\right)$, $(L^{\mathrm{pe}})^*$ *as defined in* (4.13), *defines an element of* $\left(\left(H^{-1}\left([0, t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4$.

Proof The proof is basically the same as in [262, Lemma A.2] and [263, Lemma A.3].

With the product rule we obtain

$$\begin{aligned}\left((L^{\mathrm{pe}})^*\left(\chi\left(\boldsymbol{G}^* * \nu\right)\right)\right)_i &= \chi\left((L^{\mathrm{pe}})^*\left(\boldsymbol{G}^* * \nu\right)\right)_i + \sum_{|\beta|\leq 1}\left(L^\beta \chi\right)\left(\nu_k * D^\beta \boldsymbol{G}_{ik}^*\right) \\ &= \chi\nu_i + \sum_{|\beta|\leq 1}\left(L^\beta \chi\right)\left(\nu_k * D^\beta \boldsymbol{G}_{ik}^*\right). \end{aligned} \tag{5.21}$$

Here, D^β, $\beta \in \mathbb{N}_0^4$, can be any of the derivatives $\partial_{x_j}, j \in \{1,2,3\}$, or ∂_t. We define D^β in a way that $\beta = (0,0,0,1)^T$ corresponds to ∂_t. For each β, L^β is the corresponding partial differential operator with constant coefficients of order $2 - |\beta|$ according to the product rule.

We already know that v_i and, thus, χv_i are elements of $\left(H^{-1}\left([0,t_{\text{end}}], C^l(\overline{\Omega})\right)\right)'$. This means we have to show that the sum on the right-hand side gives also an element of $\left(H^{-1}\left([0,t_{\text{end}}], C^l(\overline{\Omega})\right)\right)'$.

For what follows in this proof, we suspend summation convention. Each summand of the form $\left(v_k * D^\beta \boldsymbol{G}_{ik}^*\right)$ has compact support because

$$\operatorname{supp}\left(v_k * D^\beta \boldsymbol{G}_{ik}^*\right) = \operatorname{supp} D^\beta \left(v_k * \boldsymbol{G}_{ik}^*\right) \subset \operatorname{supp}\left(v_k * \boldsymbol{G}_{ik}^*\right) . \tag{5.22}$$

We denote $f = D^\beta \boldsymbol{G}_{ik}^*$ for any fixed combination of β, i, and k.

Remember that $\check{\chi}(x,t) = \chi(-x,-t)$ in this context.

As $\chi \in C_0^\infty(\overline{\Omega} \times [0,t_{\text{end}}])$, we have $v_k * \check{\chi} \in C_0^\infty(\overline{\Omega} \times [0,t_{\text{end}}])$ and $(v_k * \check{\chi})^\vee \in C_0^\infty(\overline{\Omega} \times [0,t_{\text{end}}])$ as well as [238, Chapter 6]

$$(v_k * \check{\chi})^\vee(x) = \langle (\mathsf{T}_x \mathsf{T}_t \check{\chi})^\vee, v_k \rangle = \langle \mathsf{T}_{-x} \mathsf{T}_{-t} \chi, v_k \rangle = v_k(\mathsf{T}_{-x} \mathsf{T}_{-t} \chi) . \tag{5.23}$$

Let $\varphi \in C_0^\infty(\mathbb{R}^3 \times \mathbb{R}_0^+)$ be such that $\varphi = 1$ in $U_1 \times [0,t_{\text{end}}]$ and $\operatorname{supp}(\varphi) \subset U_2 \times [0,t_{\text{end}}]$ with $\Omega \subset U_1 \subset U_2 \subset \hat{\Omega}$, $\operatorname{dist}(U_2, \hat{\Gamma}) > 0$. Then we obtain for every $\chi \in C_0^\infty(\overline{\Omega} \times [0,t_{\text{end}}])$

$$\begin{aligned} \langle \chi, f * v_k \rangle &= \langle \varphi\chi, f * v_k \rangle = \langle (v_k * (\varphi\chi)^\vee)^\vee, f \rangle = \langle v_k(\mathsf{T}_{-x} \mathsf{T}_{-t} (\varphi\chi)), f \rangle \\ &= \int_0^{t_{\text{end}}} \int_{\mathbb{R}^n} v_k(\mathsf{T}_{-x} \mathsf{T}_{-t} (\varphi\chi)) f(x,t) \, \mathrm{d}V(x) \mathrm{d}t . \end{aligned} \tag{5.24}$$

The last integral is well defined since the support of $v_k(\mathsf{T}_{-x} \mathsf{T}_{-t} (\varphi\chi))$ is given by

$$\begin{aligned} K &= \operatorname{supp}(v_k(\mathsf{T}_{-x} \mathsf{T}_{-t} (\varphi\chi))) = \operatorname{supp}(\varphi) - \operatorname{supp}(v_k) \\ &= \{(x-y, t-\tau) : (x,t) \in \operatorname{supp}(\varphi), (y,\tau) \in \operatorname{supp}(v_k)\} \end{aligned} \tag{5.25}$$

which is a compact subset of $U_2 \times [0,t_{\text{end}}]$ and $v(\mathsf{T}_{-x} \mathsf{T}_{-t} (\varphi\chi)) \in C_0^\infty(\overline{\Omega} \times [0,t_{\text{end}}])$. Therefore,

$$\begin{aligned} |\langle \chi, f * v_k \rangle| &= \left| \iiint_K \int v_k(\mathsf{T}_{-x} \mathsf{T}_{-t} (\varphi\chi)) f(x,t) \, \mathrm{d}V(x) \mathrm{d}t \right| \\ &\leq \left(\iiint_K \int |f(x,t)| \, \mathrm{d}V(x) \mathrm{d}t \right) \times \sup_{(x,t) \in K} |v_k(\mathsf{T}_{-x} \mathsf{T}_{-t} (\varphi\chi))| \end{aligned}$$

$$\begin{aligned}
&\leq C \left\| \nu_k \right\|_{\left(H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)\right)'} \left\| \varphi\chi \right\|_{H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)} \\
&\leq \tilde{C} \left\| \nu_k \right\|_{\left(H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)\right)'} \left\| \chi \right\|_{H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)} \\
&\leq c \left\| \chi \right\|_{H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)}
\end{aligned} \tag{5.26}$$

with c being the product of $\iiiint_K |f(x,t)| \, \mathrm{d}V(x)\mathrm{d}t$, $\left\| \nu_k \right\|_{\left(H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)\right)'}$, and $\sup\left(\left\| \varphi\chi \right\|_{H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)} : \left\| \chi \right\|_{H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)} = 1 \right)$. We used a quadruple integral here to indicate that $K \subset \mathbb{R}^4$ and we integrate over x and t with $(x,t) \in K$. Since $\iiiint_K |f(x,t)| \, \mathrm{d}V(x)\mathrm{d}t$ is finite according to Corollary 5.5, $f * \nu_k$ defines an element of $\left(H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)\right)'$. Thus, $(L^{\text{pe}})^*\left(\chi \left(\boldsymbol{G}^* * \nu\right)\right)$ defines an element of $\left(\left(H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)\right)'\right)^4$ as it is a finite sum of elements of this space.

We are now prepared to proof an analogue to Lemma 5.2 that is valid in poroelasticity.

Lemma 5.7 *Let Ω be an open bounded subset of $\mathbb{R}^3$ with the segment property and $\nu \in \left(\left(H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)\right)'\right)^4$ with compact support. Let $\boldsymbol{G}^*$ be as in Lemma 5.6. Moreover, assume $\operatorname{supp}\left(\boldsymbol{G}^* * \nu\right)$ is a compact subset of $\overline{\Omega} \times [0,t_{\text{end}}]$. Then there exists a sequence $\{\vartheta_k\}_{k\in\mathbb{N}} \subset \left(\left(C_0^\infty\left(\mathbb{R}^3 \times \mathbb{R}_0^+\right)\right)'\right)^4$ with $\operatorname{supp}(\vartheta_k) \subset \Omega \times [0,t_{\text{end}}]$ and $\left\{(L^{\text{pe}})^* \, \vartheta_k\right\}_{k\in\mathbb{N}} \subset \left(\left(H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)\right)'\right)^4$ such that $\left\{(L^{\text{pe}})^* \, \vartheta_k\right\}_{k\in\mathbb{N}}$ converges to ν in the weak-$*$ sense of $\left(\left(H^{-1}\left([0,t_{\text{end}}],C^l(\overline{\Omega})\right)\right)'\right)^4$.*

Proof As Ω has the segment property, so does $\Omega \times (0,t_{\text{end}})$. We mark the following paragraph as a quote, albeit not a literal quote, because it is only a slightly modified merging of the analogue statements in [262, page 1430], [263, page 631], and [264, page 335].

> Since $\Omega \times (0,t_{\text{end}})$ has the segment property, for every (x,t) on the boundary of $\Omega \times (0,t_{\text{end}})$, there exists a non-zero vector $\xi_{x,t} \in \mathbb{R}^4$ and an open neighborhood $U_{x,t}$ of (x,t), such that, if $(y,\tau) \in U_{x,t} \cap \left(\overline{\Omega} \times [0,t_{\text{end}}]\right)$, then $(y,\tau) + \varepsilon\xi_{x,t} \in \Omega \times (0,t_{\text{end}})$ for every $\varepsilon \in (0,1)$. Since the boundary of $\Omega \times (0,t_{\text{end}})$ is compact in $\mathbb{R}^4$, there exists a finite collection of such neighborhoods $\{U_j\}_{j=1}^J$ covering this boundary. This collection becomes an open cover of $\overline{\Omega} \times [0,t_{\text{end}}]$ by adding a suitable open set U_0 which can be chosen in a way that $\overline{U_0} \subset \Omega \times (0,t_{\text{end}})$. We denote by ξ_j the non-zero vector corresponding to U_j, when $j = 1,\ldots,J$ and set $\xi_0 = 0$. We have $\mathsf{T}_{\varepsilon\xi_j}\left[U_j \cap \left(\overline{\Omega} \times [0,t_{\text{end}}]\right)\right] \subset \Omega \times (0,t_{\text{end}})$.
>
> Let $\{\chi_j\}_{j=0}^J$ be an infinitely differentiable partition of unity corresponding to the covering $\{U_j\}_{j=0}^J$ of $\overline{\Omega} \times [0,t_{\text{end}}]$, i.e., $\chi_j \in C_0^\infty(U_j)$ and $\sum_{j=0}^J \chi_j(x,t) = 1$ for every $(x,t) \in \overline{\Omega} \times [0,t_{\text{end}}]$.

We define

$$\vartheta_j = \chi_j \left(\boldsymbol{G}^* * \nu\right) \ , \quad \vartheta_{j,\varepsilon} = \mathsf{T}_{\varepsilon\xi_j} \, \vartheta_j \ , \tag{5.27a}$$

$$\nu_j = (L^{\mathrm{pe}})^*\,\vartheta_j\,, \qquad \nu_{j,\varepsilon} = \mathsf{T}_{\varepsilon\xi_j}\,\nu_j = (L^{\mathrm{pe}})^*\,\vartheta_{j,\varepsilon}\,, \tag{5.27b}$$

with $\varepsilon \in (0,1)$. Obviously, $\operatorname{supp}(\nu_j) \subset \operatorname{supp}(\vartheta_j) \subset \overline{U}_j \cap \left(\overline{\Omega} \times [0, t_{\mathrm{end}}]\right)$ and $\operatorname{supp}(\nu_{j,\varepsilon}) \subset \operatorname{supp}(\vartheta_{j,\varepsilon}) \subset \Omega \times (0, t_{\mathrm{end}})$.

Due to Lemma 5.6, we know that $\nu_j \in \left(\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4$. Thus, $\{\nu_j\}_{j=0}^J$ gives a decomposition of ν according to

$$\sum_{j=0}^{J} \nu_j = \sum_{j=0}^{J} (L^{\mathrm{pe}})^*\left(\chi_j \vartheta\right) = (L^{\mathrm{pe}})^*\left(\left(\sum_{j=0}^{J} \chi_j\right)\vartheta\right) = (L^{\mathrm{pe}})^*\,\vartheta = \nu\,. \tag{5.28}$$

Analogously, we define

$$\nu_\varepsilon = \sum_{j=0}^{J} \nu_{j,\varepsilon} = \sum_{j=0}^{J} (L^{\mathrm{pe}})^*\,\vartheta_{j,\varepsilon} = (L^{\mathrm{pe}})^*\left(\sum_{j=0}^{J} \vartheta_{j,\varepsilon}\right) = (L^{\mathrm{pe}})^*\,\vartheta_\varepsilon \tag{5.29}$$

for which we have $\operatorname{supp}(\nu_\varepsilon) \subset \operatorname{supp}(\vartheta_\varepsilon) \subset \Omega \times (0, t_{\mathrm{end}})$. If we can show that ν_ε converges to ν in the weak-$*$ sense of $\left(\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4$ as ε approaches 0, the lemma is proven with $\{\vartheta_k\}_{k\in\mathbb{N}} = \{\vartheta_\varepsilon : \varepsilon = k^{-1},\, k \in \mathbb{N}\}$.

It suffices to show that $\nu_{j,\varepsilon}$ converges to ν_j.

To see that $\nu_{j,\varepsilon} \in \left(\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4$, we observe that $\mathsf{T}_{\varepsilon\xi_j}$ is an automorphism on $\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'$ for all $\varepsilon \in [0,1)$. For $\omega \in \left(C_0^\infty(\mathbb{R}^4)\right)^4$, we have per definition $\mathsf{T}_{\varepsilon\xi_j}\nu(\omega) = \nu(\mathsf{T}_{-\varepsilon\xi_j}\omega)$ and the set of restrictions of functions ω in $\left(C_0^\infty(\mathbb{R}^4)\right)^4$ to $\overline{\Omega} \times [0, t_{\mathrm{end}}]$ is a dense subset of $\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)^4$. Thus, $\nu_{j,\varepsilon} \in \left(\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4$ because

$$\begin{aligned}
&\left\|\nu_{j,\varepsilon}\right\|_{\left(\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4} \\
&= \sup\left(\frac{\left|\nu\left(\mathsf{T}_{-\varepsilon\xi_j}\eta\right)\right|}{\left\|\mathsf{T}_{-\varepsilon\xi_j}\eta\right\|_{\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)^4}} : \eta \in \left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)^4, \eta \neq 0\right) \\
&= \sup\left(\frac{|\nu(\eta)|}{\|\eta\|_{\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)^4}} : \eta \in \left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)^4, \eta \neq 0\right) \\
&= \left\|\nu_j\right\|_{\left(\left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4}\,.
\end{aligned} \tag{5.30}$$

For $\eta \in \left(H^{-1}\left([0,t_{\mathrm{end}}], C^l(\overline{\Omega})\right)\right)^4$ and $\omega \in \left(C_0^\infty(\mathbb{R}^4)\right)^4$, we have

$$\begin{aligned}
\left|\nu_{j,\varepsilon}(\eta) - \nu_j(\eta)\right| &\le \left|\nu_{j,\varepsilon}(\eta) - \nu_{j,\varepsilon}(\omega)\right| + \left|\nu_{j,\varepsilon}(\omega) - \nu_j(\omega)\right| + \left|\nu_j(\omega) - \nu_j(\eta)\right| \\
&= \left|\nu_{j,\varepsilon}(\eta - \omega)\right| + \left|\nu_j(\mathsf{T}_{-\varepsilon\xi_j}\omega - \omega)\right| + \left|\nu_j(\omega - \eta)\right|\,.
\end{aligned} \tag{5.31}$$

The first and last term of this estimate can become arbitrarily small if ω is chosen accordingly close to η. The second term tends to zero as ε approaches 0, because $\mathsf{T}_{-\varepsilon\xi_j}\omega$ converges to ω with $\varepsilon \to 0$ with respect to the norm in $\left(C_0^\infty(\mathbb{R}^4)\right)^4$ and, thus, certainly with respect to the norm in $\left(H^{-1}\left([0,t_{\text{end}}], C^l(\overline{\Omega})\right)\right)^4$. Consequently, we obtain weak-$*$ convergence of $v_{j,\varepsilon}$ to v_j in $\left(\left(H^{-1}\left([0,t_{\text{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4$.

In order to prove a density result, we need some more properties of solutions of the QEP (4.1) and their adjoint Eqs. (4.12).

Lemma 5.8 *If (u,p) is a strong solution of (4.1) such that $p \in C^1\left((0,t_{\text{end}}), C^2(\Omega)\right)$, $u \in C^1\left((0,t_{\text{end}}), \mathcal{C}^3(\Omega)\right)$, then $u(\cdot,t)$ and $p(\cdot,t)$ are analytic for every fixed $t \in (0,t_{\text{end}})$.*

Proof We have a look at the equations in the form (4.20) expressed with ζ instead of p. Then ζ is a solution of the heat equation and, because of our assumptions on u and p, $\zeta \in C^1\left((0,t_{\text{end}}), C^2(\Omega)\right)$. According to Mikhailov [199, Theorem 1, page 347], ζ is analytic with respect to x for every fixed $t \in (0,t_{\text{end}})$.

u now satisfies the Cauchy-Navier equation which is elliptic with right-hand side $-\frac{\alpha}{c_0\mu}\nabla_x\zeta$. Due to Hörmander [132, Theorem 9.5.1], it follows that u is analytic with respect to x for every fixed $t \in (0,t_{\text{end}})$. The same is true for $\nabla_x \cdot u$. Thus, p is analytic with respect to x for every fixed $t \in (0,t_{\text{end}})$ as a linear combination of two functions with this property.

Lemma 5.9 *If (u,p) is a solution of (4.1) with the properties as in Lemma 5.8, solutions (v,q) of (4.12) can be constructed as*

$$v^{(1)}(x,t) = (\partial_t u)(x, t_{\text{end}} - t)\,, \qquad v^{(2)}(x,t) = u(x, t_{\text{end}} - t)\,, \tag{5.32a}$$

$$q^{(1)}(x,t) = p(x, t_{\text{end}} - t)\,, \qquad q^{(2)}(x,t) = \int_0^{t_{\text{end}}-t} p(x,\eta)\,\mathrm{d}\eta\,. \tag{5.32b}$$

The same can be done vice versa, i.e., solutions (u,p) of (4.1) may be constructed from a sufficiently regular solution (v,q) of (4.12).

Proof Simply plug $(v^{(1)}, p^{(1)})$ and $(v^{(2)}, p^{(2)})$, respectively, into (4.12).

We are now ready to prove a density result for linear combinations of fundamental solutions (see (5.11)).

Theorem 5.10 *Let $\Omega \subset \mathbb{R}^3$ be an open bounded domain with the segment property and $\hat{\Omega}$ be an open bounded domain embracing Ω with (smooth) boundary $\hat{\Gamma} = \partial\hat{\Omega}$. Let $\mathcal{Y}$ be the set*

$$\mathcal{Y} = \Big\{ u \in H^{-1}\left((0,t_{\text{end}}), \mathcal{C}^l(\Omega)\right), p \in H^{-1}\left((0,t_{\text{end}}), C^l(\Omega)\right) : \\ L^{\text{pe}}(u,p) = 0, \quad p(x,0) = 0 = u(x,0)\ \forall x \in \Omega \Big\}\,.$$

Then the set of finite linear combinations of fundamental solutions with singularities (y, τ) *in* $\hat{\Gamma} \times (0, t_{\text{end}})$ *as given by* (5.11) *restricted to* $\overline{\Omega} \times [0, t_{\text{end}}]$*, which we denote by* $\mathcal{X}$*, is a dense subset of*

$$\mathcal{Y} \cap \left(H^{-1}\left([0, t_{\text{end}}], C^l(\overline{\Omega})\right)\right)^4$$

with respect to the norm in $\left(H^{-1}\left([0, t_{\text{end}}], C^l(\overline{\Omega})\right)\right)^4$.

Proof Clearly, we have $\mathcal{X} \subset \mathcal{Y} \subset \left(H^{-1}\left([0, t_{\text{end}}], C^l(\overline{\Omega})\right)\right)^4$. We want to use a duality argument based on the Hahn-Banach Theorem. Thus, we have to show that if we have $\nu \in \left(\left(H^{-1}\left([0, t_{\text{end}}], C^l(\overline{\Omega})\right)\right)'\right)^4$ such that $\nu(u, p) = 0$ for every $(u, p) \in \mathcal{X}$, then $\nu(u, p) = 0$ for every $(u, p) \in \mathcal{Y}$. For such a linear functional ν, we consider

$$\begin{aligned}\vartheta_i(y, \tau) &= \left(\boldsymbol{G}^* * \nu\right)_i (y, \tau) = \left(\nu_k * \boldsymbol{G}^*_{ik}\right)(y, \tau) \\ &= \nu_k(\mathsf{T}_y \mathsf{T}_\tau \boldsymbol{G}_{ki}) = \left(\nu(\mathsf{T}_y \mathsf{T}_\tau \boldsymbol{G})\right)_i . \end{aligned} \tag{5.33}$$

As ν can be interpreted as a distribution with compact support in $\overline{\Omega} \times [0, t_{\text{end}}]$, ϑ defines a distribution on $\mathbb{R}^3 \times \mathbb{R}_0^+$ which satisfies

$$(L^{\text{pe}})^*(\vartheta) = \nu \tag{5.34a}$$

and, thus,

$$(L^{\text{pe}})^*(\vartheta) = 0 \quad \text{in } (\mathbb{R}^3 \setminus \overline{\Omega}) \times (0, t_{\text{end}}) . \tag{5.34b}$$

Moreover, $\vartheta(y, \tau) = 0$ for all $(y, \tau) \in \hat{\Gamma} \times (0, t_{\text{end}})$ because of the properties of ν.

Let $\tilde{U}$ be a bounded connected component of $\mathbb{R}^3 \setminus \overline{\hat{\Omega}}$. According to Lemma 5.9, we can construct solutions of (4.1) from ϑ. As these solutions vanish on the boundary of $\tilde{U}$ for every $t \in (0, t_{\text{end}})$, we get from the existence and uniqueness Theorem 3.14 that they vanish on $\tilde{U} \times (0, t_{\text{end}})$. As they are analytic according to Lemma 5.8, they vanish on $U \times (0, t_{\text{end}})$, where U is the connected component of $\mathbb{R}^3 \setminus \overline{\Omega}$ such that $\tilde{U} \subset U$. Therefore, we know that $\vartheta_i(y, t_{\text{end}} - \tau)$, $i \in \{1, 2, 3\}$, and $(\partial_\tau \vartheta_4)(y, t_{\text{end}} - \tau)$ as well as $\int_0^{t_{\text{end}}-\tau} \vartheta_i(y, \eta)\, d\eta$ and $\vartheta_4(y, t_{\text{end}} - \tau)$ vanish on $U \times (0, t_{\text{end}})$. Consequently, ϑ vanishes on $U \times (0, t_{\text{end}})$.

For the unbounded connected component $\tilde{V}$ of $\mathbb{R}^3 \setminus \overline{\hat{\Omega}}$ and the corresponding unbounded connected component V of $\mathbb{R}^3 \setminus \overline{\Omega}$, we need a more sophisticated argument.

What follows is easier to understand and simpler in notation if we assume that there are functions $\nu_i(x, t)$, $i \in \{1, 2, 3, 4\}$, such that

$$\left(\nu(\mathsf{T}_y \mathsf{T}_\tau \boldsymbol{G})\right)_i = \int_0^{t_{\text{end}}} \int_{\mathbb{R}^3} \nu_k(x, t) \boldsymbol{G}_{ik}(x - y, t - \tau)\, \mathrm{d}V(x) \mathrm{d}t .$$

With this, we obtain

$$
\begin{aligned}
&\vartheta_i(y,t_{\text{end}}-\tau)\\
&=\int_0^{t_{\text{end}}}\int_{\mathbb{R}^3}\Big(\nu_1(x,t)\boldsymbol{u}^{\text{Fi}}_{1i}(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_2(x,t)\boldsymbol{u}^{\text{Fi}}_{2i}(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_3(x,t)\boldsymbol{u}^{\text{Fi}}_{3i}(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_4(x,t)p^{\text{Fi}}_i(x-y,t-(t_{\text{end}}-\tau))\Big)\,\mathrm{d}V(x)\mathrm{d}t\\
&=\int_0^{t_{\text{end}}}\int_{\mathbb{R}^3}\Big(\nu_1(x,t)\boldsymbol{u}^{\text{Fi}}_{i1}(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_2(x,t)\boldsymbol{u}^{\text{Fi}}_{i2}(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_3(x,t)\boldsymbol{u}^{\text{Fi}}_{i3}(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_4(x,t)\left(\partial_\tau u^{\text{Si}}_i\right)(x-y,t-(t_{\text{end}}-\tau))\Big)\,\mathrm{d}V(x)\mathrm{d}t\,,
\end{aligned}
\tag{5.35a}
$$

$$
\begin{aligned}
&(\partial_\tau\vartheta_4)(y,t_{\text{end}}-\tau)\\
&=\int_0^{t_{\text{end}}}\int_{\mathbb{R}^3}\Big(\nu_1(x,t)\left(\partial_\tau u^{\text{Si}}_1\right)(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_2(x,t)\left(\partial_\tau u^{\text{Si}}_2\right)(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_3(x,t)\left(\partial_\tau u^{\text{Si}}_3\right)(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_4(x,t)\left(\partial_\tau p^{\text{Si}}\right)(x-y,t-(t_{\text{end}}-\tau))\Big)\,\mathrm{d}V(x)\mathrm{d}t\\
&=\partial_\tau\int_0^{t_{\text{end}}}\int_{\mathbb{R}^3}\Big(\nu_1(x,t)u^{\text{Si}}_1(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_2(x,t)u^{\text{Si}}_2(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_3(x,t)u^{\text{Si}}_3(x-y,t-(t_{\text{end}}-\tau))\\
&\qquad+\nu_4(x,t)p^{\text{Si}}(x-y,t-(t_{\text{end}}-\tau))\Big)\,\mathrm{d}V(x)\mathrm{d}t\,,
\end{aligned}
\tag{5.35b}
$$

where $i \in \{1,2,3\}$ in this case. Considering $u_i(y,\tau) = \vartheta_i(y, t_{\text{end}} - \tau)$, $i \in \{1,2,3\}$, and $p(y,\tau) = (\partial_\tau \vartheta_4)(y, t_{\text{end}} - \tau)$ yields for $\tau > 0$

$$
\begin{aligned}
&\nabla_y \cdot u(y,\tau) \\
&= \int_0^{t_{\text{end}}} \int_{\mathbb{R}^3} \Big(\nu_1(x,t) \nabla_y \cdot \boldsymbol{u}^{\text{Fi},1}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_2(x,t) \nabla_y \cdot \boldsymbol{u}^{\text{Fi},2}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_3(x,t) \nabla_y \cdot \boldsymbol{u}^{\text{Fi},3}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_4(x,t) \nabla_y \cdot \left(\partial_\tau u_1^{\text{Si}}\right)(x-y, t-(t_{\text{end}}-\tau)) \Big) \, \mathrm{d}V(x) \mathrm{d}t \\
&= -\frac{C_1}{C_2} \int_0^{t_{\text{end}}} \int_{\mathbb{R}^3} \Big(\nu_1(x,t) p_1^{\text{Fi}}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_2(x,t) p_2^{\text{Fi}}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_3(x,t) p_3^{\text{Fi}}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_4(x,t) \left(\partial_\tau p^{\text{Si}}\right)(x-y, t-(t_{\text{end}}-\tau)) \Big) \, \mathrm{d}V(x) \mathrm{d}t \\
&= -\frac{C_1}{C_2} \int_0^{t_{\text{end}}} \int_{\mathbb{R}^3} \Big(\nu_1(x,t) \left(\partial_\tau u_1^{\text{Si}}\right)(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_2(x,t) \left(\partial_\tau u_2^{\text{Si}}\right)(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_3(x,t) \left(\partial_\tau u_3^{\text{Si}}\right)(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_4(x,t) \left(\partial_\tau p^{\text{Si}}\right)(x-y, t-(t_{\text{end}}-\tau)) \Big) \, \mathrm{d}V(x) \mathrm{d}t \\
&= -\frac{C_1}{C_2} \partial_\tau \int_0^{t_{\text{end}}} \int_{\mathbb{R}^3} \Big(\nu_1(x,t) u_1^{\text{Si}}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_2(x,t) u_2^{\text{Si}}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_3(x,t) u_3^{\text{Si}}(x-y, t-(t_{\text{end}}-\tau)) \\
&\qquad + \nu_4(x,t) p^{\text{Si}}(x-y, t-(t_{\text{end}}-\tau)) \Big) \, \mathrm{d}V(x) \mathrm{d}t \,, \qquad (5.36)
\end{aligned}
$$

where $\boldsymbol{u}^{\mathrm{Fi},1}$ means the first column of $\boldsymbol{u}^{\mathrm{Fi}}$ and so on, and C_1, C_2 as defined in (4.26). As $\tau \in (0, t_{\mathrm{end}})$, we also have $t_{\mathrm{end}} - \tau \in (0, t_{\mathrm{end}})$ and, thus,

$$(\partial_\tau \vartheta_4)\,|_{\hat{\Gamma}\times(0,t_{\mathrm{end}})} = 0 = \big(\nabla_y \cdot u\big)\,|_{\hat{\Gamma}\times(0,t_{\mathrm{end}})}. \tag{5.37}$$

Moreover, the corresponding ζ vanishes on $\hat{\Gamma} \times (0, t_{\mathrm{end}})$. According to the estimates in Lemma 5.4, $\partial_\tau \vartheta_4$, u, and ζ also vanish as $\|y\| \to \infty$. As ζ satisfies the heat equation, the maximum principle for solutions of parabolic differential equations, [101, Theorem 1, page 34], holds. Therefore, ζ has to vanish on $\tilde{V} \times (0, t_{\mathrm{end}})$. As a consequence, u satisfies the Cauchy-Navier Eq. (4.20a) with vanishing right-hand side. Hence, the same argument as in [264] leads to the result that u vanishes on $\tilde{V}$ for every $\tau \in (0, t_{\mathrm{end}})$, since the estimates used in [264] also hold in our case. As both, ζ and u are analytic on the larger domain $V \times (0, t_{\mathrm{end}})$, they also have to vanish on this domain. Due to the fact that $(\partial_\tau \vartheta_4)$ is a linear combination of ζ and u, it also vanishes on $V \times (0, t_{\mathrm{end}})$. The same chain of reasoning holds for $u_i(y, \tau) = \int_0^{t_{\mathrm{end}}-\tau} \vartheta(y, \eta)\, \mathrm{d}\eta$, $i \in \{1, 2, 3\}$, and $p(y, \tau) = \vartheta_4(y, t_{\mathrm{end}} - \tau)$. Consequently, ϑ vanishes on $V \times (0, t_{\mathrm{end}})$.

Putting the results for the bounded and unbounded connected components of $\mathbb{R}^3 \setminus \overline{\Omega}$ together, we get that the support of ϑ is a compact subset of $\Omega \times (0, t_{\mathrm{end}})$. This means that Lemma 5.7 provides a sequence $(\vartheta_k)_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \big((L^{\mathrm{pe}})^*\, \vartheta_k\big)(u, p) = \nu(u, p) \quad \text{for every } (u, p) \in \mathcal{Y}\,. \tag{5.38}$$

To prove the theorem, it now suffices to show that $\big\{(L^{\mathrm{pe}})^*\, \vartheta_k\big\}\,(u, p) = 0$ for every $(u, p) \in \mathcal{Y}$. To this end, choose $\chi \in C_0^\infty(\mathbb{R}^3 \times [0, t_{\mathrm{end}}])$ such that χ is equal to 1 in a neighborhood of $\mathrm{supp}(\vartheta_k)$. Then we have $(u, p) = (\chi u, \chi p)$ on this neighborhood and, since u and p are real analytic functions on $\Omega \times (0, t_{\mathrm{end}})$, we get $(\chi u, \chi p) \in \mathcal{C}_0^\infty(\mathbb{R}^3 \times (0, t_{\mathrm{end}})) \times C_0^\infty(\mathbb{R}^3 \times (0, t_{\mathrm{end}}))$. Further on, $(\chi u, \chi p)$ and (u, p) as well as the corresponding partial derivatives coincide on $\Omega \times (0, t_{\mathrm{end}})$. This yields

$$\begin{aligned}\big((L^{\mathrm{pe}})^*\, \vartheta_k\big)(u, p) &= \big((L^{\mathrm{pe}})^*\, \vartheta_k\big)(\chi u, \chi p) = \big\langle (\chi u, \chi p), (L^{\mathrm{pe}})^*\, \vartheta_k \big\rangle \\ &= \langle L^{\mathrm{pe}}(\chi u, \chi p), \vartheta_k \rangle = \langle L^{\mathrm{pe}}(u, p), \vartheta_k \rangle = 0\,. \end{aligned} \tag{5.39}$$

Here $\langle \cdot, \cdot \rangle$ is the dual pairing of $\mathcal{C}_0^\infty\big(\mathbb{R}^3 \times (0, t_{\mathrm{end}})\big) \times C_0^\infty\big(\mathbb{R}^3 \times (0, t_{\mathrm{end}})\big)$ with its dual space of distributions. As this equation holds for all $k \in \mathbb{N}$, it also holds in the limit $k \to \infty$ and we finally have shown that

$$\nu(u, p) = 0 \text{ for every } (u, p) \in \mathcal{Y}\,. \tag{5.40}$$

This concludes the proof of Theorem 5.10.

Remark 5.11 Analogous versions of Lemmata 5.6, 5.7, and Theorem 5.10 hold in case of the heat equation. The corresponding space of functions can be improved to be $C^l(\hat{\Omega} \times [0, t_{\mathrm{end}}])$. This is due to the fact that

$$\lim_{t\to 0} G^{\mathrm{Heat}}(x - y, t) = \delta(x - y), \tag{5.41}$$

thus, vanishing on $\overline{\Omega}$ for $y \in \hat{\Gamma}$. The proofs can be done mainly in the same way as above and are, therefore, omitted. The density results for the heat equation obtained that way are equivalent to the ones given by Kupradze [163] and Johansson et al. [143].

5.2.2 *Density Results for Non-vanishing Initial Conditions*

As we have seen in the derivation of Theorem 3.14, prescribing ζ at $t = 0$ is a suitable initial condition. There are different ways how to incorporate non-vanishing initial values of ζ.

The first idea is to extend Theorem 5.10 by extending the set of singularity points which we consider. With a fixed time $\hat{t} > 0$, we say that instead of $(y, \tau) \in \hat{\Gamma} \times (0, t_{\text{end}})$, we allow $(y, \tau) \in (\hat{\Gamma} \times (-\hat{t}, t_{\text{end}})) \cup (\hat{\Omega} \times \{-\hat{t}\})$. Except for more technicalities in notation, the proofs of Theorem 5.10 and the preceding lemmata stay the same and the method is extended to contain non-vanishing initial values. This is consistent to the fact that prescribing initial values of ζ is a suitable initial condition, as we have seen in the derivation of Theorem 3.14 that at $t = 0$, $\zeta(\cdot, 0)$ can be determined from a given pressure $p(\cdot, 0)$ and prescribed boundary conditions for the displacement vector $u(\cdot, 0)$. It may seem as if we lose the benefit of dimension reduction because we need to include source points in the domain $\hat{\Omega}$ and collocation points in Ω. However, this is not the case, because we need such points only for one specific point in time. The domain on which we find a solution is the four-dimensional set $\Omega \times (0, t_{\text{end}})$, whereas the domain on which we are working is three-dimensional, either with three dimensions in space and only one point in time or with two dimensions in space and one time dimension.

Another idea is related to the fact that ζ is a solution of the heat Eq. (4.20b). In a two-dimensional setting, Johannson et al. [143] show that the set of fundamental solutions with singularities in $\hat{\Gamma} \times (-t_{\text{end}}, t_{\text{end}})$ is dense in $L^2(\Gamma \times (-t_{\text{end}}, t_{\text{end}}))$ and $L^2(\Omega)$. In this article, it is required that the domain Ω is bounded by a C^2-smooth bounding surface Γ. This is necessary in order to use jump and limit relations of layer potentials and Green's identities for the heat equation. Since we are only interested in density in $L^2(\Omega)$, we may ease this restriction. It suffices if Γ is a connected Lipschitz boundary.

Theorem 5.12 *The set of fundamental solutions of the heat equation* (4.20b) $\left\{G^{\text{Heat}}(\cdot - y_{(m)}, \cdot - \tau_n)\right\}_{n,m\in\mathbb{N}}$, $\tau_n < 0$, *with* $(y_{(m)}, \tau_n)_{n,m\in\mathbb{N}}$ *a countable dense set of points in* $\hat{\Gamma} \times (-t_{\text{end}}, t_{\text{end}})$, *form a linearly independent and dense set in* $L^2(\Omega)$.

Proof The proof follows exactly the one of [143, Theorem 3.2] and is included here for the convenience of the reader.

First we show linear independence. For $\tau_n < 0$, let

$$\zeta(x,t) = \sum_{n=1}^{N}\sum_{m=1}^{M} c_{nm} G^{\text{Heat}}(x - y_{(m)}, t - \tau_n) \,, \tag{5.42}$$

$$\zeta(x,0) = \sum_{n=1}^{N}\sum_{m=1}^{M} c_{nm} G^{\text{Heat}}(x - y_{(m)}, 0 - \tau_n) \,,$$

with $\zeta(x,0) = 0$ for all $x \in \Omega$. We have to show that $c_{nm} = 0$ for all $n, m \in \mathbb{N}$. As $\tau_n < 0$, $G^{\text{Heat}}(\cdot - y_{(m)}, t - \tau_n)$ is real analytic for all $t \in (0, t_{\text{end}})$, see Lemma 4.8. Thus, $\zeta(x,0) = 0$ for all $x \in \mathbb{R}^3$ and $\zeta \in C^2(\mathbb{R}^3 \times [0, t_{\text{end}}])$. Moreover, ζ satisfies the heat equation in $\mathbb{R}^3 \times [0, t_{\text{end}}]$ and by Lemma 5.4 we also have

$$|\zeta(x,t)| \leq B \exp\Big(\beta \, \|x\|^2\Big) \tag{5.43}$$

for some constant $B, \beta \in \mathbb{R}^+$. From the Corollary to Theorem 3[3] in [144], we get that $\zeta(x,t) = 0$ in $\mathbb{R}^3 \times [0, t_{\text{end}}]$. According to [244], this yields $\zeta(x,t) = 0$ in $\mathbb{R}^3 \times [-t_{\text{end}}, t_{\text{end}}]$.

Assume that (x,t) approaches $(y_{(m_0)}, \tau_{n_0})$ in a way that the ratio

$$\frac{\|x - y_{m_0}\|^2}{4C_2(t - \tau_{n_0})} \tag{5.44}$$

remains bounded. Then $c_{n_0 m_0} G^{\text{Heat}}(x - y_{(m_0)}, t - \tau_{n_0})$ may become arbitrarily large while all other summands in (5.42) remain bounded. This is in contradiction to $\zeta(x,t) = 0$ in $\mathbb{R}^3 \times [-t_{\text{end}}, t_{\text{end}}]$. Therefore, $c_{n_0 m_0} = 0$ and linear independence is established.

To show density, we assume that there is a function $f \in L^2(\Omega)$ such that

$$\int_{\Omega} f(x) G^{\text{Heat}}(x - y_{(m)}, t - \tau_n) \, \mathrm{d}V(x) = 0 \quad \forall n, m \in \mathbb{N} \,. \tag{5.45}$$

This is obviously true for all functions $f \in L^2(\Omega)$ if $\tau_n \geq 0$, since $G^{\text{Heat}}(x - y_{(m)}, t - \tau_n)|_{x \in \Omega} = 0$ for those τ_n. Thus, if we can show density, we have shown density for the set considered in Theorem 5.12.

Let $w(x,t)$ be a solution of the heat equation in $\Omega \times (0, t_{\text{end}})$ with initial condition $w(x,0) = w^{(0)}(x)$ and boundary condition $w(x,t) = 0$ for $(x,t) \in \Gamma \times (0, t_{\text{end}})$. From Green's identities for the heat equation (see, e.g., [57, Proposition 2.19]), we obtain

[3] There are two theorems numbered as 3 in [144]. We refer here to the second one.

$$\int_0^{t_{\text{end}}} \int_{\Gamma} \left(\nabla_x w(x,t) \cdot n(x)\right) G^{\text{Heat}}(x - y_{(m)}, t - \tau_n) \, \mathrm{d}S(x) \mathrm{d}t = 0 \quad \forall n, m \in \mathbb{N} \, . \tag{5.46}$$

From Remark 5.11, it follows that $(\nabla_x w(x,t) \cdot n(x)) = 0$ on $\Gamma \times (0, t_{\text{end}})$, since the considered set of fundamental solutions is dense in $L^2(\Gamma \times (0, t_{\text{end}}))$ (see also [163]). As w is a solution of the heat equation which vanishes on the boundary for all times and with vanishing normal derivative on the boundary for all times, w must vanish on the whole domain $\Omega \times (0, t_{\text{end}})$. Again, the results of [244] give us that $w(x,t) = 0$ on $\Omega \times (-t_{\text{end}}, t_{\text{end}})$. This means that $w(x,0) = 0$ for all $x \in \Omega$ and, therefore, $w^{(0)} \equiv 0$.

The most interesting feature of Theorem 5.12 is that it reduces the computational resources we need. If ζ is given at $t = 0$, we only need one scalar fundamental solution for each source point instead of the full set of 16 entries in the 4×4-tensor of fundamental solutions. This stays true if an initial pressure $p(\cdot, 0)$ is given as long as we are not explicitly interested in the displacement field for $t = 0$.

The considerations of this section give rise to the idea of an alternative approach on how to use fundamental solutions in a numerical solution scheme for quasistatic poroelasticity.

5.2.3 *Alternative Aspects to the Method of Fundamental Solutions*

As mentioned above, the previous section gives rise to the idea of an alternative MFS. For this purpose, we consider (4.20) with vanishing right-hand sides, i.e., $f \equiv 0, h \equiv 0$, such that

$$-\frac{c_0(\lambda + \mu) + \alpha^2}{c_0 \mu} \nabla_x \left(\nabla_x \cdot u\right) - \nabla_x^2 u = -\frac{\alpha}{c_0 \mu} \nabla_x \zeta \, , \tag{5.47a}$$

$$\partial_t \zeta - C_2 \nabla_x^2 \zeta = 0 \tag{5.47b}$$

with C_2 as given in (4.26). In this formulation, we have the homogeneous heat equation (5.47b) for ζ and Eq. (5.47a) can be treated as an inhomogeneous Cauchy-Navier equation for u. The corresponding fundamental solution tensor is given by

$$\boldsymbol{G}^{\text{alt}}(x,t) = \begin{pmatrix} \boldsymbol{u}^{\text{CN}}(x)\delta(t) & u^{\text{Si}}(x,t) \\ 0 & G^{\text{Heat}}(x,t) \end{pmatrix} \, . \tag{5.48}$$

As Eqs. (5.47) are decoupled, we can treat them separately. Density proofs for the heat equation are given in [143, 163]. Moreover, an alternative proof is implicitly included in Sect. 5.2.1. Density of fundamental solutions of the Cauchy-Navier

equation is proven in [264]. Here, a proof of density of fundamental solutions for the system of equations of static thermoelasticity is also included, which differs from Eqs. (5.47) by incorporating the Laplace equation instead of the heat equation. This proof follows the concepts of Sect. 5.2.1 and is reduced to the proof that for a suitable annihilating functional ν, the fourth entry of $\left(\boldsymbol{G}^{\mathrm{alt}}\right)^T * \nu$ vanishes on $\mathbb{R}^3 \setminus \overline{\Omega}$. As the fourth columns of $\boldsymbol{G}^{\mathrm{alt}}$ and $\boldsymbol{G}$ differ only by a constant factor in front of their fourth entry which contains the heat kernel, this is already part of the proof in Sect. 5.2.1. Therefore, we can conclude that for every $t > 0$, the set of finite linear combinations of fundamental solutions $\boldsymbol{G}^{\mathrm{alt}}$ with singularities (y, τ) in $\hat{\Gamma} \times (0, t_{\mathrm{end}})$ restricted to $\overline{\Omega} \times [0, t_{\mathrm{end}}]$ is dense in

$$\mathcal{Y}^{\mathrm{alt}}(t) \cap \left(C^l(\overline{\Omega})\right)^4$$

with

$$\mathcal{Y}^{\mathrm{alt}}(t) = \left\{u(\cdot, t) \in \mathcal{C}^2(\Omega), \zeta(\cdot, t) \in C^2(\Omega) : (u, \zeta) \text{ solution to Eqs. (5.47)}\right\}$$

with respect to the norm in $\left(C^l(\overline{\Omega})\right)^4$. As a consequence, the corresponding set of tensors of solutions of $L^{\mathrm{pe}}(u, p) = 0$ is dense in the corresponding solution space of the QEP. This basically means that we can omit the fi-parts of the fundamental solutions in Ansatz (5.10) and, thus, further reduce the computational demands with respect to memory.

Remark 5.13 Considering the validity of this alternative ansatz and the fact that the fundamental solutions of fi-type, $\boldsymbol{u}^{\mathrm{fi}}$ and p^{fi}, can be interpreted as the gradient of the fundamental solutions of Si-type, u^{Si} and p^{Si}, Ansatz (5.10) may be interpreted as a partial multipole series of first order (cf. [80]), where a full first order multipole series would also involve partial first order derivatives of $\boldsymbol{u}^{\mathrm{CN}}$ and p^{St}. To decide whether this interpretation is justified, further research on the theoretical properties of Ansatz (5.10) is needed, which is out of the scope of this thesis.

Chapter 6
Numerical Results

As Chap. 5 introduced the MFS and gave theoretical results, the results on the numerical performance of the method are given in this chapter. First, we briefly discuss implementation aspects of the method with regard to the considered exemplary problems, like choice of collocation or source points in space and time, underlying structure of matrices to the resulting systems of linear equations and chosen solution methods for them. Next, we perform a parameter study of the method. For this purpose, we consider some initial boundary value problems on the square $(-1,1)^2$, i.e., a two-dimensional domain. The corresponding dimensioned domain would be $(-x_0,x_0)^2$ with the characteristic lengthscale x_0.

For well performing parameters, we will compare results using a robust solution method (singular value decomposition) to such obtained by a simple Gaussian least-squares method as well as results obtained by a reduced Ansatz (6.5) to the full Ansatz (5.10). Furthermore, we investigate another, alternative way how the MFS may be used for time-dependent problems in order to decrease memory requirements. Finally, we discuss solutions with steep gradients and give an insight into applications of the MFS on the three-dimensional cube $(-1,1)^3$.

All results presented in this chapter were obtained by algorithms implemented in MATLAB® [190], which was also used for graphical illustrations.

6.1 Implementation Details

We use the reduced ansatz of Sect. 5.2.3 without the fi-parts of the fundamental solutions taking account of the results of Sect. 5.2.2 for the choice of source points with singularity at negative times. Figure 6.1 illustrates the location of collocation and source points in space and time for initial boundary value problems in quasistatic poroelasticity. The domain Ω under consideration is the square

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4_6

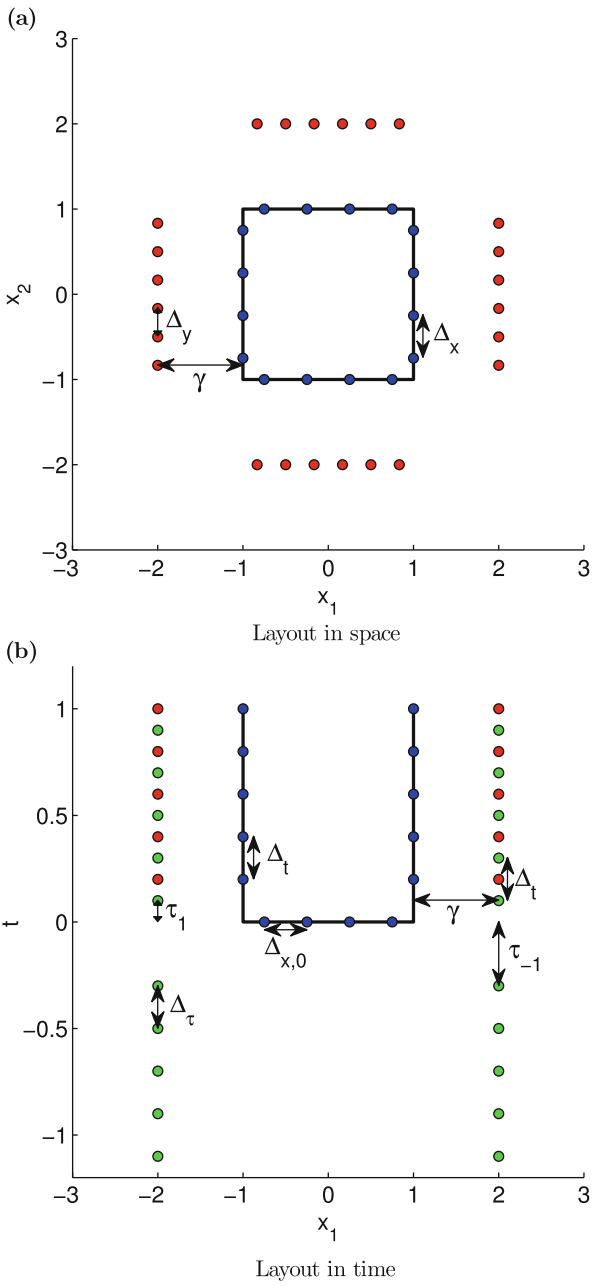

Fig. 6.1 Positions of the collocation points (*blue*) and source points (*red* for CN- and St-parts, *green* for Si-parts) used to compute a solution to the QEP (4.1) with a reduced MFS. For the meaning of the different marked quantities, see Table 6.1

Table 6.1 Parameters of the method, see also Fig. 6.1

Symbol	Denotes
γ	Distance between Γ and $\hat{\Gamma}$
Δ_x	Spatial distance between neighboring collocation points at Γ
$\Delta_{x,0}$	Spatial distance between neighboring collocation points in Ω for $t = 0$
Δ_y	Spatial distance between neighboring source points at $\hat{\Gamma}$
Δ_t	Temporal distance between collocation points at Γ and source points at $\hat{\Gamma}$ for positive τ
Δ_τ	Temporal distance between source points at $\hat{\Gamma}$ for negative τ
τ_1	Smallest positive source point time for Si-parts
τ_{-1}	Absolute value of the largest negative source point time for Si-parts

$(-1, 1)^2$, which is a bounded open domain with the segment property as required, e.g., in Theorem 5.10. Here, blue dots correspond to the collocation points, red dots to the source points associated with the CN- and St-parts of the fundamental solutions, respectively, and green dots to the source points associated with the Si-parts of the fundamental solutions. The meaning of the parameters is explained in Table 6.1. Please note that we always choose $0 < \tau_1 < \Delta_t$.

For simplicity, we assume $\{2\Delta_x^{-1}, 2\Delta_{x,0}^{-1}, 2\Delta_y^{-1}\} \subset \mathbb{N}$. Moreover, we define throughout this chapter

$$I = \left\lfloor \frac{2}{\Delta_x} \right\rfloor, \quad I_{\text{int}} = \left\lfloor \frac{2}{\Delta_{x,0}} \right\rfloor, \quad J = \left\lfloor \frac{t_{\text{end}}}{\Delta_t} \right\rfloor, \quad M = \left\lfloor \frac{2}{\Delta_y} \right\rfloor, \tag{6.1}$$

with the usual floor function $\lfloor \cdot \rfloor$ as defined in Chap. 2. The choice of source points is in correspondence with results presented in [8] which suggest that the MFS shows good performance if the source points $y_{(m)}$ are chosen by choosing points $y_{(m,\Gamma)}$ on the actual boundary of Ω and translate them in direction of the normal to the boundary, i.e.

$$y_{(m)} = y_{(m,\Gamma)} + \gamma n(y_{(m,\Gamma)}) \tag{6.2}$$

with $\gamma \in \mathbb{R}^+$ and $n(y_{(m,\Gamma)})$ being the normal to the boundary of Ω in $y_{(m,\Gamma)}$.[1] Subscripts in parentheses are used to refer to the numbering of spatial coordinates of source points or collocation points to avoid confusion with the components of spatial coordinates. In distinction from [8], we do not compute approximate normal vectors at each point $y_{(m,\Gamma)}$. The lengths of those approximate normal vectors depend on the distance between neighboring points that varies in [8]. As neighboring collocation and source points have always the same distance in our examples, we use the exactly known unit normal vector at each point. Please note that neither any collocation

[1] It is usually not necessary to define the corresponding pseudo-boundary $\hat{\Gamma}$ explicitly.

points nor source points are positioned at the corners of the square (Fig. 6.1a) as it is unclear which boundary condition should be implemented if different boundary conditions are prescribed on neighboring edges and there is no well-defined normal vector at a corner. Such a kind of normal distribution of source points was also used in [288].

Besides of choosing a spatial location for the singularity of the fundamental solutions at a distance from Γ, the source points for the Si-parts of the fundamental solutions also have to be shifted in time, as $p^{\mathrm{Si}}(x-y, t-\tau)$ approaches $\delta(x-y)$ as t approaches τ. We do this by choosing a positive $\tau_1 < \Delta_t$ as the smallest positive τ for our source points and distributing them in time with the same minimal distance Δ_t as the collocation points. The spatial coordinates of the singularities of fundamental solutions of CN- and St-parts on the one hand and Si-parts on the other hand are the same. For source points with singularity at negative points in time, it is mentioned in [143] that it may not be necessary to distribute them in the whole interval $[-t_{\mathrm{end}}, 0]$, whereas it is shown in [131] that choosing all of those source points close to $t = 0$ results in large approximation errors. In our approach, the minimal distance in time between the collocation points at $t = 0$ and the source points with negative τ is given by τ_{-1}, whereas the largest is given by $\tau_{-1} + (J-1)\Delta_\tau$. This means that source points with negative τ are positioned in the time interval $[-(\tau_{-1} + (J-1)\Delta_\tau), -\tau_{-1}]$.

In accordance with [143], the collocation points in the domain Ω at $t = 0$ are chosen such that the number of collocation points in $\Omega \times \{0\}$ is approximately equal to the number of collocation points in $\Gamma \times (0, t_{\mathrm{end}}]$. This yields the definition of $\Delta_{x,0}$ such that $I_{\mathrm{int}} = \left\lfloor \sqrt{I \cdot J} \right\rfloor$. Hence, $\Delta_{x,0}$ depends on Δ_x, Δ_t and the considered t_{end}.

The set of collocation points for the initial conditions can be written as

$$\begin{aligned} &\left\{x_1 : x_1 = -1 + \left(i - \frac{1}{2}\right)\Delta_{x,0}, i \in \mathbb{N}, i \le I_{\mathrm{int}}\right\} \\ \times &\left\{x_2 : x_2 = -1 + \left(j - \frac{1}{2}\right)\Delta_{x,0}, j \in \mathbb{N}, j \le I_{\mathrm{int}}\right\} \\ \times &\{t = 0\}\,, \end{aligned} \tag{6.3a}$$

and for the boundary conditions as

$$\begin{aligned} \Bigg(&\left\{x = \left(-1, -1 + \left(i - \frac{1}{2}\right)\Delta_x\right)^T, i \in \mathbb{N}, i \le I\right\} \\ \cup &\left\{x = \left(1, -1 + \left(i - \frac{1}{2}\right)\Delta_x\right)^T, i \in \mathbb{N}, i \le I\right\} \\ \cup &\left\{x = \left(-1 + \left(i - \frac{1}{2}\right)\Delta_x, -1\right)^T, i \in \mathbb{N}, i \le I\right\} \end{aligned}$$

$$\cup \left\{ x = \left(-1 + \left(i - \frac{1}{2} \right) \Delta_x, 1 \right)^T , i \in \mathbb{N}, i \le I \right\} \Bigg)$$
$$\times \{ t = j\Delta_t, j \in \mathbb{N}, j \le J \} \ . \tag{6.3b}$$

Similarly, we have three sets of source points. For the CN- and St-parts, we have

$$\Bigg(\left\{ y = \left(-1 - \gamma, -1 + \left(m - \frac{1}{2} \right) \Delta_y \right)^T , m \in \mathbb{N}, m \le M \right\}$$
$$\cup \left\{ y = \left(1 + \gamma, -1 + \left(m - \frac{1}{2} \right) \Delta_y \right)^T , m \in \mathbb{N}, m \le M \right\}$$
$$\cup \left\{ y = \left(-1 + \left(m - \frac{1}{2} \right) \Delta_y, -1 - \gamma \right)^T , m \in \mathbb{N}, m \le M \right\}$$
$$\cup \left\{ y = \left(-1 + \left(m - \frac{1}{2} \right) \Delta_y, 1 + \gamma \right)^T , m \in \mathbb{N}, m \le M \right\} \Bigg)$$
$$\times \{ \tau = n\Delta_t, n \in \mathbb{N}, n \le J \} \ . \tag{6.4a}$$

The time coordinates given here are just to provide a better orientation although CN- and St-parts of the fundamental solutions as used here are time-independent and the time-dependence of the solution to a given initial boundary value problem is included in the coefficients associated with those parts (see (5.10) and (6.5)). For the Si-parts with singularity at positive times τ, we have the source points

$$\Bigg(\left\{ y = \left(-1 - \gamma, -1 + \left(m - \frac{1}{2} \right) \Delta_y \right)^T , m \in \mathbb{N}, m \le M \right\}$$
$$\cup \left\{ y = \left(1 + \gamma, -1 + \left(m - \frac{1}{2} \right) \Delta_y \right)^T , m \in \mathbb{N}, m \le M \right\}$$
$$\cup \left\{ y = \left(-1 + \left(m - \frac{1}{2} \right) \Delta_y, -1 - \gamma \right)^T , m \in \mathbb{N}, m \le M \right\}$$
$$\cup \left\{ y = \left(-1 + \left(m - \frac{1}{2} \right) \Delta_y, 1 + \gamma \right)^T , m \in \mathbb{N}, m \le M \right\} \Bigg)$$
$$\times \{ \tau = \tau_1 + (n-1)\Delta_t, n \in \mathbb{N}, n \le J \} \ , \tag{6.4b}$$

and for those with negative τ

$$\begin{aligned}
&\Bigg(\quad \left\{ y = \left(-1-\gamma, -1+\left(m-\frac{1}{2}\right)\Delta_y \right)^T, m \in \mathbb{N}, m \le M \right\} \\
&\quad \cup \left\{ y = \left(1+\gamma, -1+\left(m-\frac{1}{2}\right)\Delta_y \right)^T, m \in \mathbb{N}, m \le M \right\} \\
&\quad \cup \left\{ y = \left(-1+\left(m-\frac{1}{2}\right)\Delta_y, -1-\gamma \right)^T, m \in \mathbb{N}, m \le M \right\} \\
&\quad \cup \left\{ y = \left(-1+\left(m-\frac{1}{2}\right)\Delta_y, 1+\gamma \right)^T, m \in \mathbb{N}, m \le M \right\} \Bigg) \\
&\times \{ \tau = -\tau_{-1} - (n-1)\Delta_\tau, n \in \mathbb{N}, n \le J \} \ . \qquad \text{(6.4c)}
\end{aligned}$$

The condition on n in (6.4c) ensures that there are as much source points with negative τ as with positive τ. The ordering of collocation and source points is based on the above description.

To achieve density of source points as required in the results of Chap. 5, taking the limits $\Delta_y \to 0$, $\Delta_\tau \to 0$, $\tau_{-1} \to 0$, $\Delta_t \to 0$, and $\tau_1 \to 0$ would not be sufficient as this would not yield a dense set in a connected pseudo-boundary. To arrive at a situation which satisfies the conditions on the distribution of source points which we used in our theoretical results, we would have to explicitly define an embracing open bounded domain $\hat{\Omega}$ with a smooth boundary $\hat{\Gamma}$. Choosing a distribution that achieves the goal of yielding a dense set on $\hat{\Gamma}$ would contradict the attempt of staying close to the results of [8] and, thus, probably lead to a poorer numerical performance. Therefore, we accept this little drawback with regards to theoretical results.

The MFS ansatz we use here can be written in the form

$$\begin{aligned}
u_k(x,t) = &\sum_{n=1}^{J} \sum_{m=1}^{4M} c^{(1)}_{(n-1)\,4M+m}\, u_k^{\mathrm{Si}}(x - y_{(m)}, t - \tau_{-n}) \\
&+ \sum_{\substack{n=1 \\ \tau_n < t}}^{J} \sum_{m=1}^{4M} c^{(2)}_{3(n-1)\,4M+3m}\, u_k^{\mathrm{Si}}(x - y_{(m)}, t - \tau_n) \\
&+ \sum_{m=1}^{4M} \sum_{l=1}^{2} \left(c(t)\right)^{(2)}_{3(n-1)\,4M+3(m-1)+l} \boldsymbol{u}_{kl}^{\mathrm{CN}}(x - y_{(m)}) \ , \qquad \text{(6.5a)} \\
p(x,t) = &\sum_{n=1}^{J} \sum_{m=1}^{4M} c^{(1)}_{(n-1)\,4M+m}\, p^{\mathrm{Si}}(x - y_{(m)}, t - \tau_{-n})
\end{aligned}$$

$$+\sum_{\substack{n=1\\ \tau_n<t}}^{J}\sum_{m=1}^{4M} c^{(2)}_{3(n-1)\,4M+3m}\, p^{\mathrm{Si}}(x-y_{(m)},t-\tau_n)$$

$$+\sum_{m=1}^{4M}\sum_{l=1}^{2}\left(c(t)\right)^{(2)}_{3(n-1)\,4M+3(m-1)+l} p^{\mathrm{St}}_{l}(x-y_{(m)})\ . \tag{6.5b}$$

The indices of source points as used here are understood as

$$\left(y_{(m)}\right)_{m=1}^{M} = \left\{ y=\left(-1-\gamma,-1+\left(m-\frac{1}{2}\right)\Delta_y\right)^T, m\in\mathbb{N}, m\le M\right\}\ , \tag{6.6a}$$

$$\left(y_{(M+m)}\right)_{m=1}^{M} = \left\{ y=\left(1+\gamma,-1+\left(m-\frac{1}{2}\right)\Delta_y\right)^T, m\in\mathbb{N}, m\le M\right\}\ , \tag{6.6b}$$

$$\left(y_{(2M+m)}\right)_{m=1}^{M} = \left\{ y=\left(-1+\left(m-\frac{1}{2}\right)\Delta_y,-1-\gamma\right)^T, m\in\mathbb{N}, m\le M\right\}\ , \tag{6.6c}$$

$$\left(y_{(3M+m)}\right)_{m=1}^{M} = \left\{ y=\left(-1+\left(m-\frac{1}{2}\right)\Delta_y,1+\gamma\right)^T, m\in\mathbb{N}, m\le M\right\}\ . \tag{6.6d}$$

Instead of distinguishing coefficients for different parts of the fundamental solutions by denoting them as a, a^{CN}, b, and $b^{(0)}$ as in (5.10), all coefficients are condensed into one coefficient vector, denoted by c, whose superscript indicates whether the source points of the corresponding parts of the fundamental solutions are located at negative τ ($c^{(1)}$) or positive τ ($c^{(2)}$). Moreover, the subscript of c indicates the exact source point of the corresponding fundamental solution part in case of Si-parts.

Assuming collocation in the points given by (6.3), an initial condition that provides $\zeta(\cdot,0)$ and Dirichlet boundary conditions (i.e., u and p given on the whole boundary), we get from Ansatz (6.5) in shorthand notation the following system of linear equations

$$\underbrace{\begin{pmatrix} \boldsymbol{A}^{(\mathrm{IC})} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{A}^{(-1)} & \boldsymbol{A}^{(1)} & \ddots & \vdots & \mathbf{0} & \mathbf{0} \\ \vdots & \boldsymbol{A}^{(2)} & \ddots & \mathbf{0} & \vdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \boldsymbol{A}^{(1)} & \mathbf{0} & \vdots \\ \boldsymbol{A}^{(-(J-1))} & \boldsymbol{A}^{(J-1)} & \cdots & \boldsymbol{A}^{(2)} & \boldsymbol{A}^{(1)} & \mathbf{0} \\ \boldsymbol{A}^{(-J)} & \boldsymbol{A}^{(J)} & \cdots & \boldsymbol{A}^{(3)} & \boldsymbol{A}^{(2)} & \boldsymbol{A}^{(1)} \end{pmatrix}}_{\boldsymbol{A}} \begin{pmatrix} c^{(1)} \\ c^{(2)} \end{pmatrix} = \begin{pmatrix} r^{(1)} \\ r^{(2)} \end{pmatrix} \tag{6.7a}$$

with the various submatrices given by

$$\begin{aligned} \boldsymbol{A}^{(\mathrm{IC})}_{(j-1)I_{\mathrm{int}}+i,(n-1)\,4M+m} =& \zeta^{\mathrm{Si}}(x_{1,(i)} - y_{1,(m)}, x_{2,(j)} - y_{2,(m)}, -\tau_{-n}), \\ & i,j,m,n \in \mathbb{N}, i,j \le I_{\mathrm{int}}, m \le 4M, n \le J\,, \end{aligned} \tag{6.7b}$$

$$\begin{aligned} \left(\boldsymbol{A}^{(-j)}_{k,(n-1)\,4M+m}\right)^{3i}_{k=3(i-1)+1} =& \begin{pmatrix} u_1^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_j - \tau_{-n}) \\ u_2^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_j - \tau_{-n}) \\ p^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_j - \tau_{-n}) \end{pmatrix}, \\ & i,m,n \in \mathbb{N}, i \le 4I, m \le 4M, n \le J\,, \end{aligned} \tag{6.7c}$$

$$\begin{aligned} &\left(\boldsymbol{A}^{(1)}_{kl}\right)^{3i,3m}_{k=3(i-1)+1,l=3(m-1)+1} = \\ &\begin{pmatrix} \boldsymbol{u}_{11}^{\mathrm{CN}}(x_{(i)} - y_{(m)}) & \boldsymbol{u}_{12}^{\mathrm{CN}}(x_{(i)} - y_{(m)}) & u_1^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_1 - \tau_1) \\ \boldsymbol{u}_{21}^{\mathrm{CN}}(x_{(i)} - y_{(m)}) & \boldsymbol{u}_{22}^{\mathrm{CN}}(x_{(i)} - y_{(m)}) & u_2^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_1 - \tau_1) \\ p_1^{\mathrm{St}}(x_{(i)} - y_{(m)}) & p_2^{\mathrm{St}}(x_{(i)} - y_{(m)}) & p^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_1 - \tau_1) \end{pmatrix}, \\ & i,m \in \mathbb{N}, i \le 4I, m \le 4M\,, \end{aligned} \tag{6.7d}$$

$$\begin{aligned} \left(\boldsymbol{A}^{(j)}_{kl}\right)^{3i,3m}_{k=3(i-1)+1,l=3(m-1)+1} =& \begin{pmatrix} 0\ 0\ u_1^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_j - \tau_1) \\ 0\ 0\ u_2^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_j - \tau_1) \\ 0\ 0\ p^{\mathrm{Si}}(x_{(i)} - y_{(m)}, t_j - \tau_1) \end{pmatrix}, \\ & i,j,m \in \mathbb{N}, i \le 4I, 2 \le j \le J, m \le 4M\,, \end{aligned} \tag{6.7e}$$

the coefficient vector $(c^{(1)}, c^{(2)})^T$ as defined by (6.5), and the vector on the right-hand side

$$r^{(1)}_{(j-1)I_{\text{int}}+i} = \zeta^0(x_{1,(i)}, x_{2,(j)}, 0), \quad i,j \in \mathbb{N}, i \leq I_{\text{int}}, j \leq I_{\text{int}} , \tag{6.7f}$$

$$r^{(2)}_{3(j-1)\,4I+3(i-1)+1} = u_1^{\text{BC}}(x_{(i)}, t_j), \qquad i,j \in \mathbb{N}, i \leq 4I, j \leq J , \tag{6.7g}$$

$$r^{(2)}_{3(j-1)\,4I+3(i-1)+2} = u_2^{\text{BC}}(x_{(i)}, t_j), \qquad i,j \in \mathbb{N}, i \leq 4I, j \leq J , \tag{6.7h}$$

$$r^{(2)}_{3(j-1)\,4I+3i} = p^{\text{BC}}(x_{(i)}, t_j), \qquad i,j \in \mathbb{N}, i \leq 4I, j \leq J , \tag{6.7i}$$

given by the initial and boundary conditions (superscript BC) evaluated at the corresponding collocation points. Please note that the different kinds of submatrices in (6.7) have different sizes and that the zero matrices in (6.7a) also differ in size. For boundary conditions involving normal stresses or normal fluid velocities, the system matrix can be constructed similarly with the corresponding fundamental stresses and fluid velocities given in Appendix A.

As seen in (6.7a), the system matrix $\boldsymbol{A}$ is a lower triangular block matrix. This is typical for all time-dependent systems independently of the types of boundary conditions when solved in the time domain because the fundamental solutions always vanish for $t - \tau < 0$. As the submatrices $\boldsymbol{A}^{(j)}$, $1 \leq j \leq J$, are repeated along the corresponding diagonals, each of them only has to be computed once, yielding an increased speed for the computation of the matrix. Similarly, memory requirements can be reduced as each $\boldsymbol{A}^{(j)}$ has to be stored only once.

It may seem natural to exploit the block structure in (6.7) in a solution scheme that utilizes a kind of blockwise forward substitution to reduce computational cost. Such a scheme can be seen as a kind of time-marching approach: First, a part of the solution is computed that satisfies the initial condition and its contribution to the boundary conditions is calculated. Then, the boundary conditions are modified by those values and another part of the solution which satisfies the modified boundary conditions for all collocation points at Δ_t is computed. Further boundary values are modified by the boundary values of this new part of the solution and more parts are computed analogously until the boundary conditions are satisfied for all collocation points. However, due to the usual poor condition of matrices in the MFS [262], such a scheme is prone to severe error amplification and propagation, leading to rapidly increasing errors. Therefore, it would probably be rather unstable. Moreover, System (6.7) is usually over- or underdetermined, depending on the choice of Δ_x and Δ_y. Thus, we can only solve each subsystem approximately in case of overdetermination or must choose one of several solutions in case of underdetermination. In each case, we may expect that the probability of bad performance due to amplification and propagation of errors increases. If we wanted to prevent approximation errors due to over- or underdetermination, we would have to choose $\Delta_x = \Delta_y$ and would have to use another way to choose collocation points at $t = 0$ to guarantee that $\boldsymbol{A}^{(\text{IC})}$ is a square matrix. Nevertheless, the condition of each submatrix would still be poor, such that error propagation also has to be expected if all submatrices are square.

An alternative approach based on fundamental solutions which implements a time-marching scheme is discussed in Sect. 6.5.

Remark 6.1 Smyrlis and Karageorghis studied especially the underdetermined systems in the context of the MFS for the Laplace equation in two dimensions [266, 267]. They suggested two different ways to choose a coefficient vector, on the one hand as a (weighted) least-squares minimizer among all possible coefficient vectors, on the other hand as a minimizer of the functional corresponding to the variational formulation associated with the differential equation. However, as the second choice incorporates integration of derivatives of fundamental solutions, we do not implement it here as we explicitly intend to develop an integration-free method.

System (6.7) is solved by using a singular-value decomposition $\boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T$ with orthonormal matrices $\boldsymbol{U}$, $\boldsymbol{V}$, and $\boldsymbol{\Sigma}$ being a diagonal matrix whose non-zero diagonal elements are the singular-values of the system matrix $\boldsymbol{A}$, i.e., the square roots of the eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ [157, Section 5.4], [268, Chapter 6], [269, Appendix A]. A solution to the system $\boldsymbol{A}c = r$ is then given by $\boldsymbol{V}\,\boldsymbol{\Sigma}^{-1}\,\boldsymbol{U}^T r$. As $\boldsymbol{V}\,\boldsymbol{\Sigma}^{-1}\,\boldsymbol{U}^T$ is the pseudo-inverse to $\boldsymbol{A}$, we obtain a least-squares solution in the under- as well as the overdetermined case [157]. Moreover, there are numerically robust algorithms to determine the singular-value decomposition [268]. We use the svd routine implemented in MATLAB, which is based on an algorithm by Golub and Reinsch [119, 203] without any further regularization.

For our numerical examples, we choose a fixed set of material parameters. In accordance with Ostermann [215, 216], we assume a reservoir made of sandstone, specifically Berea sandstone, for which the material parameters are given, e.g., by Schanz in [247]. The parameters relevant for us are listed in Table 6.2. For comparison, we also give the parameters for soil [247]. Please note that Schanz gives porosity ϕ and a material parameter R instead of c_0. They are related by

$$c_0 = \frac{\phi}{\mathrm{R}}. \tag{6.8}$$

For Berea sandstone, a characteristic time of $t_0 = 2.5\ \mathrm{h} = 9{,}000\ \mathrm{s}$ corresponds to a characteristic length of $x_0 \approx 101.29\ \mathrm{m}$, whereas $t_0 = 2.5\ \mathrm{d} = 60\ \mathrm{h} = 216{,}000\ \mathrm{s}$

Table 6.2 Material parameters as given by Schanz [247]. Note that the set of parameters used in [247] differs from the one used within this thesis. We give here the parameters which are relevant for us

		Berea sandstone	Soil
λ	$[\mathrm{N/m^2}]$	$4.0 \cdot 10^9$	$1.4467 \cdot 10^8$
μ	$[\mathrm{N/m^2}]$	$6.0 \cdot 10^9$	$9.8 \cdot 10^7$
c_0	$[\mathrm{m^2/N}]$	$7.6809 \cdot 10^{-11}$	$1.92 \cdot 10^{-10}$
α	[1]	0.867	0.981
k	$[\mathrm{m^4/Ns}]$	$1.9 \cdot 10^{-10}$	$3.55 \cdot 10^{-9}$

corresponds to $x_0 \approx 496.23$ m and $t_0 = 10$ d $= 240$ h $= 864{,}000$ s corresponds to $x_0 \approx 992.45$ m, according to (3.47). We compare this to soil by calculating the characteristic times for these characteristic lengths and get $t_0 \approx 29{,}491$ s ≈ 8.2 h for $x_0 = 101.29$ m, $t_0 \approx 707{,}790$ s ≈ 196.6 h ≈ 8.2 d for $x_0 = 496.23$ m and $t_0 \approx 2{,}831{,}158$ s ≈ 786.4 h ≈ 32.8 d for $x_0 = 992.45$ m. Thus, the characteristic times for soil are roughly three times longer than for Berea sandstone.

The dimensionless parameters C_1, C_2 (see (4.26)), C_3, and C_4 (see (4.35)) for Berea sandstone are

$$C_1 \approx 0.4337, \quad C_2 \approx 1.3464, \quad C_3 \approx 0.6163, \quad C_4 \approx 0.6225. \tag{6.9}$$

In the following, we use several kinds of relative errors which are either rooted mean square errors (rms) or maximal errors (max). We usually employ relative errors, indicated by the superscript rel, which are defined as

$$\mathrm{E}^{\mathrm{rel}}(x, t) = \frac{\left|p^{\mathrm{num}}(x_i, t_j) - p^{\mathrm{ana}}(x_i, t_j)\right|}{\left|p^{\mathrm{ana}}(x_i, t_j)\right|}, \tag{6.10a}$$

$$\mathrm{E}^{\mathrm{rel}}_{\mathrm{max}} = \max_{1 \le i \le I_{\mathrm{tot}}} \max_{1 \le j \le J} \frac{\left|p^{\mathrm{num}}(x_i, t_j) - p^{\mathrm{ana}}(x_i, t_j)\right|}{\left|p^{\mathrm{ana}}(x_i, t_j)\right|}, \tag{6.10b}$$

$$\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}(x) = \left(\frac{\sum_{j=1}^{J} \left|p^{\mathrm{num}}(x, t_j) - p^{\mathrm{ana}}(x, t_j)\right|^2}{\sum_{j=1}^{J} \left|p^{\mathrm{ana}}(x, t_j)\right|^2} \right)^{\frac{1}{2}}, \tag{6.10c}$$

$$\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}(t) = \left(\frac{\sum_{i=1}^{I_{\mathrm{tot}}} \left|p^{\mathrm{num}}(x_i, t) - p^{\mathrm{ana}}(x_i, t)\right|^2}{\sum_{i=1}^{I_{\mathrm{tot}}} \left|p^{\mathrm{ana}}(x_i, t)\right|^2} \right)^{\frac{1}{2}}, \tag{6.10d}$$

for the pressure p and similarly for the components of the displacement vector u, poroelastic stress tensor $\boldsymbol{\sigma}^{\mathrm{pe}}$, and fluid velocity v_f. The superscripts num and ana denote the numerical approximate solution and the exact analytical solution, respectively. I_{tot} is the total number of points x_i under consideration which is in general neither identical to the total number of collocation points on the boundary for a fixed time t_j, given by $4I$, nor the number of collocation points for $t = 0$, which is I_{int}^2.

6.2 Parameter Studies on the Square $(-1, 1)^2$

The parameters of our MFS, as given in Table 6.1, can be organized in three groups:

- Spatial parameters, consisting of Δ_x, Δ_y, and γ,
- Temporal parameters for negative times, consisting of τ_{-1} and Δ_τ, and
- Temporal parameters for positive times, consisting of τ_1 and Δ_t.

We investigate each of these groups separately, fixing all parameters that belong to the other two sets. For this purpose, we consider the following exemplary initial boundary value problems:

Example 6.2 Let $\Omega = (-1,1)^2$, $\Gamma = \partial\Omega$.

Approximate $u^{\mathrm{Si}}(x_1-2,x_2-2,t+1)$, $p^{\mathrm{Si}}(x_1-2,x_2-2,t+1)$ in $\Omega\times(0,t_{\mathrm{end}})$, given the initial condition $\zeta^0(x_1,x_2,0) = \zeta^{\mathrm{Si}}(x_1-2,x_2-2,1)$ on Ω and Dirichlet boundary conditions, i.e., $u^{\mathrm{BC}}(x_1,x_2,t) = u^{\mathrm{Si}}(x_1-2,x_2-2,t+1)$, $p^{\mathrm{BC}}(x_1,x_2,t) = p^{\mathrm{Si}}(x_1-2,x_2-2,t+1)$ on $\Gamma\times(0,t_{\mathrm{end}})$.

Example 6.3 Let $\Omega = (-1,1)^2$, $\Gamma = \partial\Omega$.

Approximate $u^{\mathrm{Si}}(x_1-2,x_2-2,t+1)$, $p^{\mathrm{Si}}(x_1-2,x_2-2,t+1)$ in $\Omega\times(0,t_{\mathrm{end}})$, given the initial condition $\zeta^0(x_1,x_2,0) = \zeta^{\mathrm{Si}}(x_1-2,x_2-2,1)$ on Ω and the mixed boundary conditions

$$
\begin{aligned}
&u^{\mathrm{BC}}(x_1,x_2,t) = u^{\mathrm{Si}}(x_1-2,x_2-2,t+1)\,,\\
&p^{\mathrm{BC}}(x_1,x_2,t) = p^{\mathrm{Si}}(x_1-2,x_2-2,t+1)\,,\\
&\text{for } x_1\in\{-1,1\}, x_2\in(-1,1), t\in(0,t_{\mathrm{end}})\,,\\
&\boldsymbol{\sigma}^{\mathrm{pe,BC}}(x_1,x_2,t)n(x) = \boldsymbol{\sigma}^{\mathrm{pe,Si}}(x_1-2,x_2-2,t+1)\begin{pmatrix}0\\ x_2\end{pmatrix}\,,\\
&v_{f,2}^{\mathrm{BC}}(x_1,x_2,t) = x_2\partial_{x_2}p^{\mathrm{Si}}(x_1-2,x_2-2,t+1)\,,\\
&\text{for } x_1\in(-1,1), x_2\in\{-1,1\}, t\in(0,t_{\mathrm{end}})\,.
\end{aligned}
$$

Example 6.4 Let $\Omega = (-1,1)^2$, $\Gamma = \partial\Omega$.

Approximate $u^{\mathrm{fi}}(x_1-2,x_2-2,t+1)$, $p^{\mathrm{fi}}(x_1-2,x_2-2,t+1)$ in $\Omega\times(0,t_{\mathrm{end}})$, given the initial condition $\zeta^0(x_1,x_2,0) = \zeta^{\mathrm{fi}}(x_1-2,x_2-2,1)$ on Ω and Dirichlet boundary conditions, i.e., $u^{\mathrm{BC}}(x_1,x_2,t) = u^{\mathrm{fi}}(x_1-2,x_2-2,t+1)$, $p^{\mathrm{BC}}(x_1,x_2,t) = p^{\mathrm{fi}}(x_1-2,x_2-2,t+1)$ on $\Gamma\times(0,t_{\mathrm{end}})$.

Example 6.5 Let $\Omega = (-1,1)^2$, $\Gamma = \partial\Omega$.

Approximate $u^{\mathrm{fi}}(x_1-2,x_2-2,t+1)$, $p^{\mathrm{fi}}(x_1-2,x_2-2,t+1)$ in $\Omega\times(0,t_{\mathrm{end}})$, given the initial condition $\zeta^0(x_1,x_2,0) = \zeta^{\mathrm{fi}}(x_1-2,x_2-2,1)$ on Ω and the mixed boundary conditions

$$
\begin{aligned}
&u^{\mathrm{BC}}(x_1,x_2,t) = u^{\mathrm{fi}}(x_1-2,x_2-2,t+1)\,,\\
&p^{\mathrm{BC}}(x_1,x_2,t) = p^{\mathrm{fi}}(x_1-2,x_2-2,t+1)\,,\\
&\text{for } x_1\in\{-1,1\}, x_2\in(-1,1), t\in(0,t_{\mathrm{end}})\,,\\
&\boldsymbol{\sigma}^{\mathrm{pe,BC}}(x_1,x_2,t)n(x) = \boldsymbol{\sigma}^{\mathrm{pe,fi}}(x_1-2,x_2-2,t+1)\begin{pmatrix}0\\ x_2\end{pmatrix}\,,\\
&v_{f,2}^{\mathrm{BC}}(x_1,x_2,t) = x_2\partial_{x_2}p^{\mathrm{fi}}(x_1-2,x_2-2,t+1)\,,\\
&\text{for } x_1\in(-1,1), x_2\in\{-1,1\}, t\in(0,t_{\mathrm{end}})\,.
\end{aligned}
$$

Please note that Examples 6.2 and 6.3 as well as Examples 6.4 and 6.5 have the same solution, respectively.

The development in time for several poroelastic quantities, evaluated at the origin $x = (0,0)$, is given in Fig. 6.2 for the solution of Examples 6.2 and 6.3. As can be seen, the pressure p (c) and diagonal elements of the stress tensor $\boldsymbol{\sigma}^{\mathrm{pe}}$ (d-e) are not monotone. This is due to the poroelastic coupling and would not be seen in simple diffusion processes (see also Remark 6.10). The time development of the pressure in Examples 6.4 and 6.5 is monotone (Fig. 6.7). However, other quantities, e.g., u_1, show a non-monotone behavior. Figures 6.3 and 6.8 show the same quantities for Examples 6.2 and 6.3 as well as Examples 6.4 and 6.5, respectively, for fixed time $t = 1$ on the whole domain. They are also depicted for $t = 5$ in Ω in Figs. 6.4 and 6.9. Please note that every subfigure uses its own scale on the ordinate or colorbar, as using the same scale for all subfigures would not allow to recognize any dependencies on x or t. The directions of the vector valued quantities u and v_f can be seen in Fig. 6.5 for $t = 1$ and in Fig. 6.6 for $t = 5$ for Examples 6.2 and 6.3. The monopole nature of the Si-fundamental solution is clearly recognizable (cf. the electric field of a point charge). The directions of u and v_f for Examples 6.4 and 6.5 are given in Fig. 6.10 for $t = 1$ and in Fig. 6.11 for $t = 5$. Differently to the Si-quantities, the directions change considerably over time.

Evaluation for each example is based on the errors defined in (6.10). The relative rooted mean square error with respect to time, $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}(x)$ is evaluated using the time grid as given by (6.3b) for each of the (spatial) points given by

$$\mathfrak{I}_1 = \{x_1 = -1 + 0.5\,i, i \in \mathbb{N}_0, i \leq 2\} \times \{x_2 = -1 + 0.5\,j, j \in \mathbb{N}_0, j \leq 2\}\,. \tag{6.11}$$

The relative rooted mean square error in the domain Ω, $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}(t)$ is evaluated at the times given by

$$\mathfrak{I}_t = \{1, 2, 3, 4, 5\}\,, \tag{6.12}$$

using the grid

$$\mathfrak{I}_2 = \{x_1 = -1 + 0.04\,i, i \in \mathbb{N}_0, i \leq 50\} \times \{x_2 = -1 + 0.04\,j, j \in \mathbb{N}_0, j \leq 50\}\,, \tag{6.13}$$

i.e., $I_{\mathrm{tot}} = 2{,}601$ in (6.10d).

6.2.1 *Varying Spatial Method Parameters*

First, we investigate the influence of the spatial parameters Δ_x, Δ_y, and γ on the performance of the MFS. Therefore, we fix $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$ in this section, yielding $J = 100$.

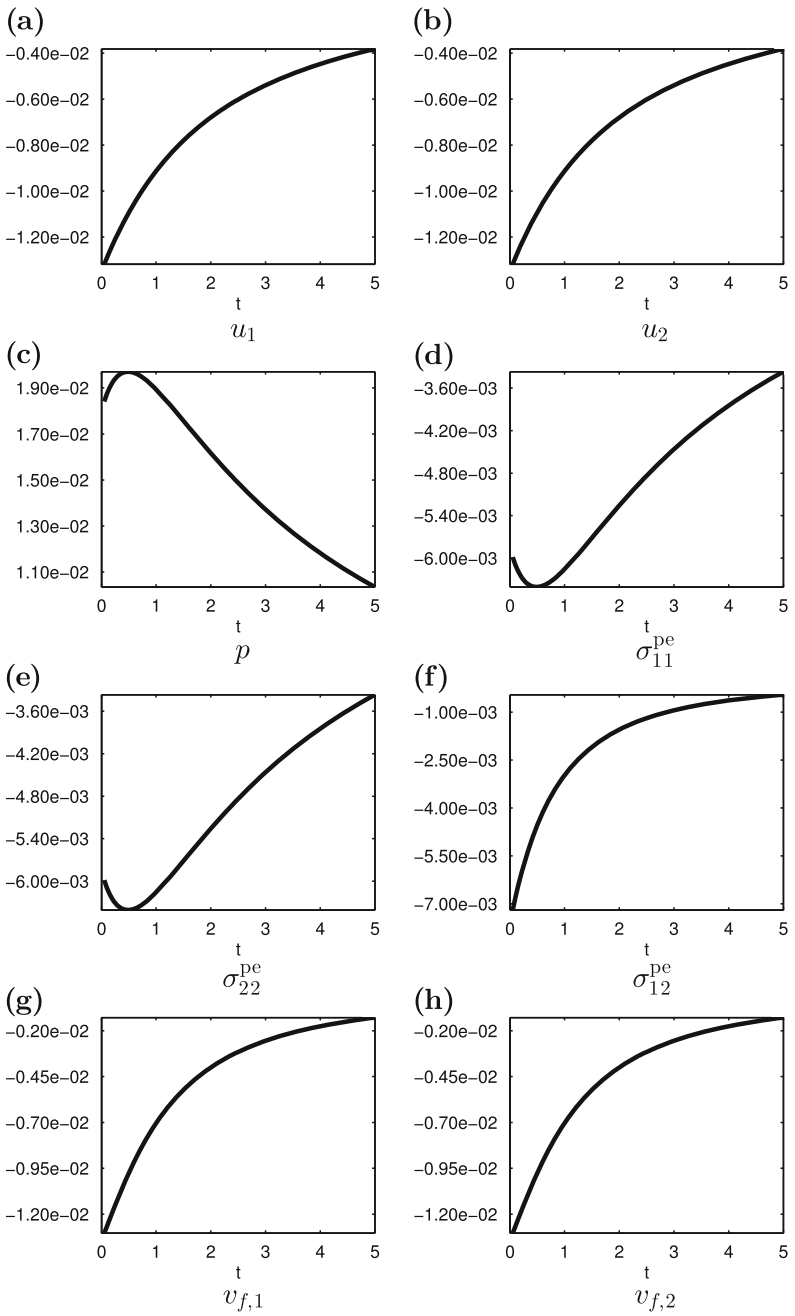

Fig. 6.2 Time development of the poroelastic quantities for Examples 6.2 and 6.3 evaluated at the origin $x = 0$. Please note that different scales are used for different quantities as indicated by the different scales on the ordinates. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

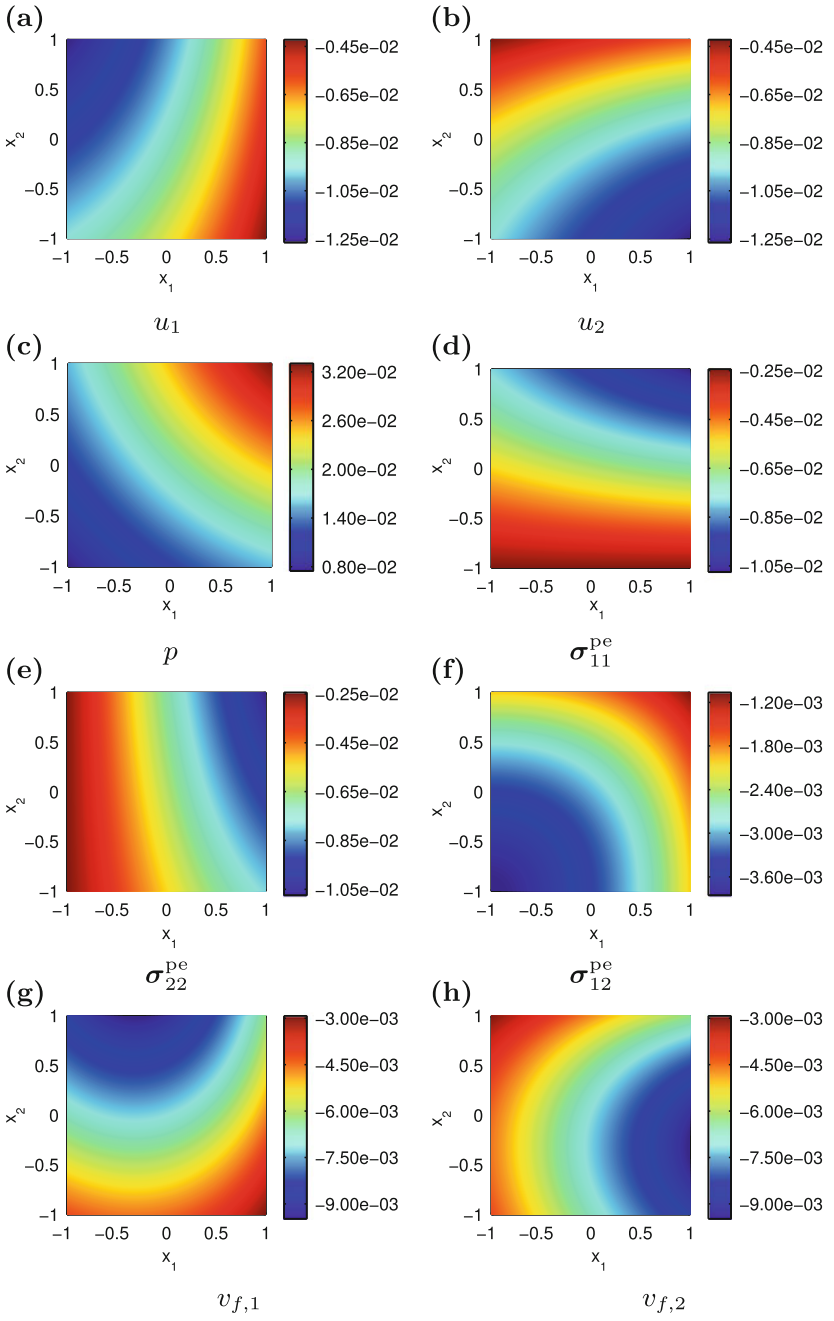

Fig. 6.3 Pseudocolor plots of the poroelastic quantities for Examples 6.2 and 6.3 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Please note that different scales are used for different quantities as indicated by the different scales on the colorbars. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

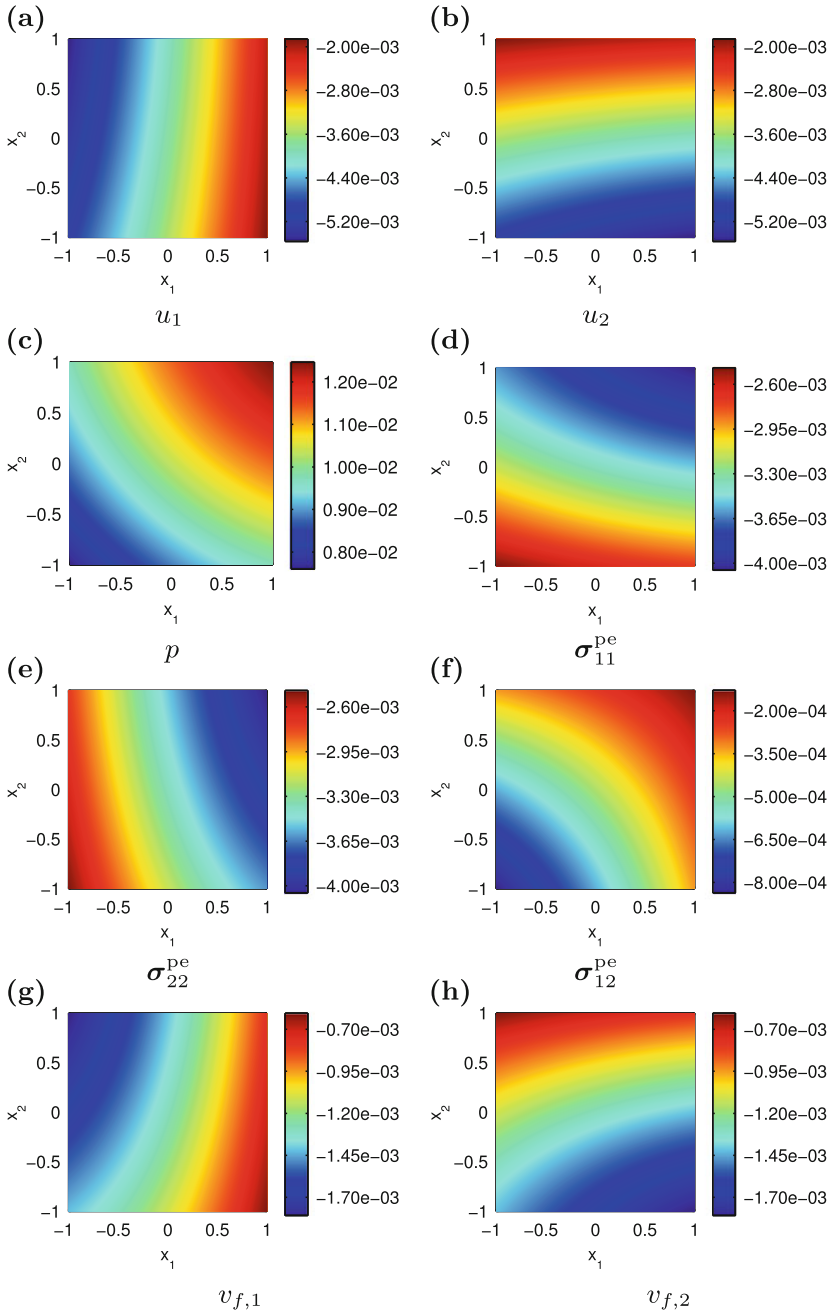

Fig. 6.4 Pseudocolor plots of the poroelastic quantities for Examples 6.2 and 6.3 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Please note that different scales are used for different quantities as indicated by the different scales on the colorbars. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

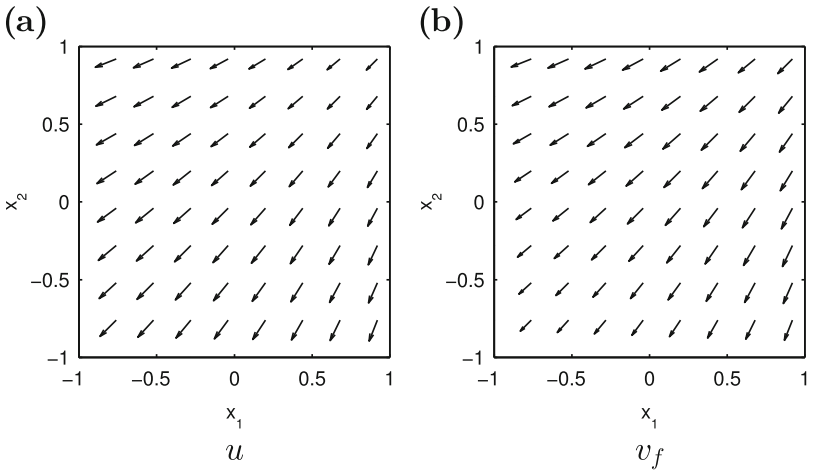

Fig. 6.5 Directions of the vector-valued poroelastic quantities u and v_f for Examples 6.2 and 6.3 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Please compare Fig. 6.3a, b and g, h as the above plots do not give sufficient information about the strength of the fields

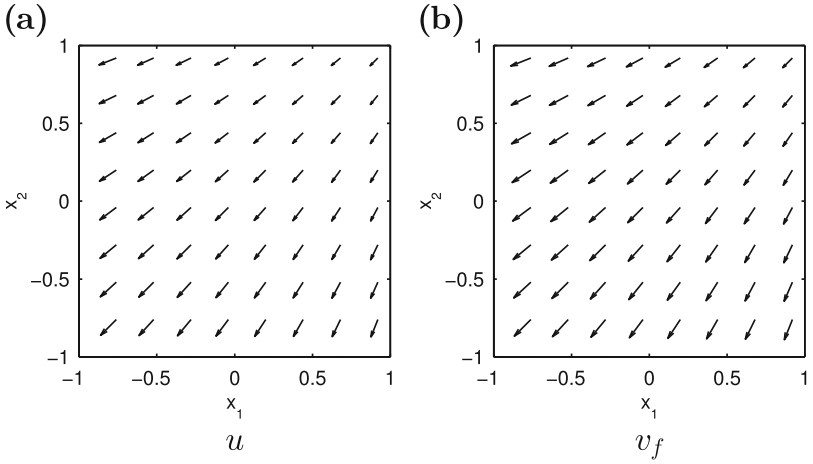

Fig. 6.6 Directions of the vector-valued poroelastic quantities u and v_f for Examples 6.2 and 6.3 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Please compare Fig. 6.4a, b and g, h as the above plots do not give sufficient information about the strength of the fields

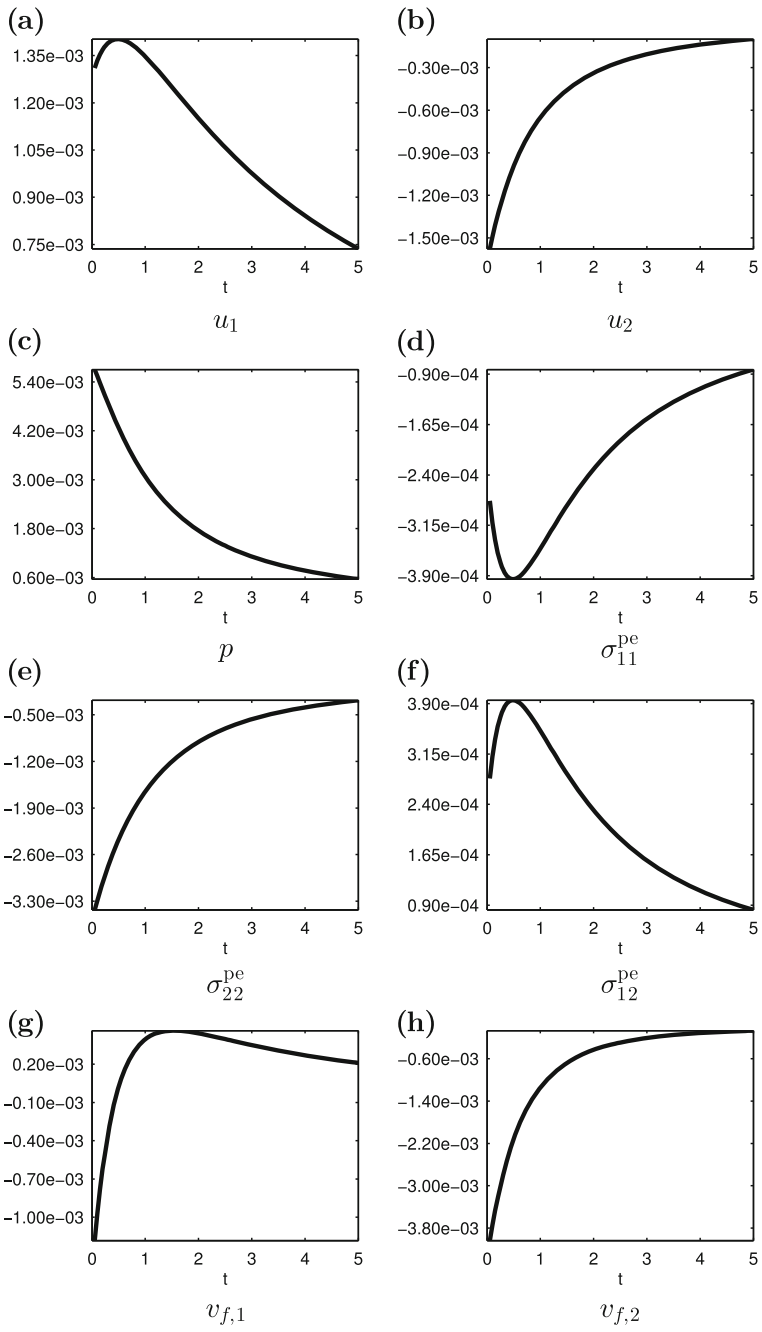

Fig. 6.7 Time development of the poroelastic quantities for Examples 6.4 and 6.5 evaluated at the origin $x = 0$. Please note that different scales with different signs are used for different quantities as indicated by the different scales on the ordinates. All quantities are dimensionless

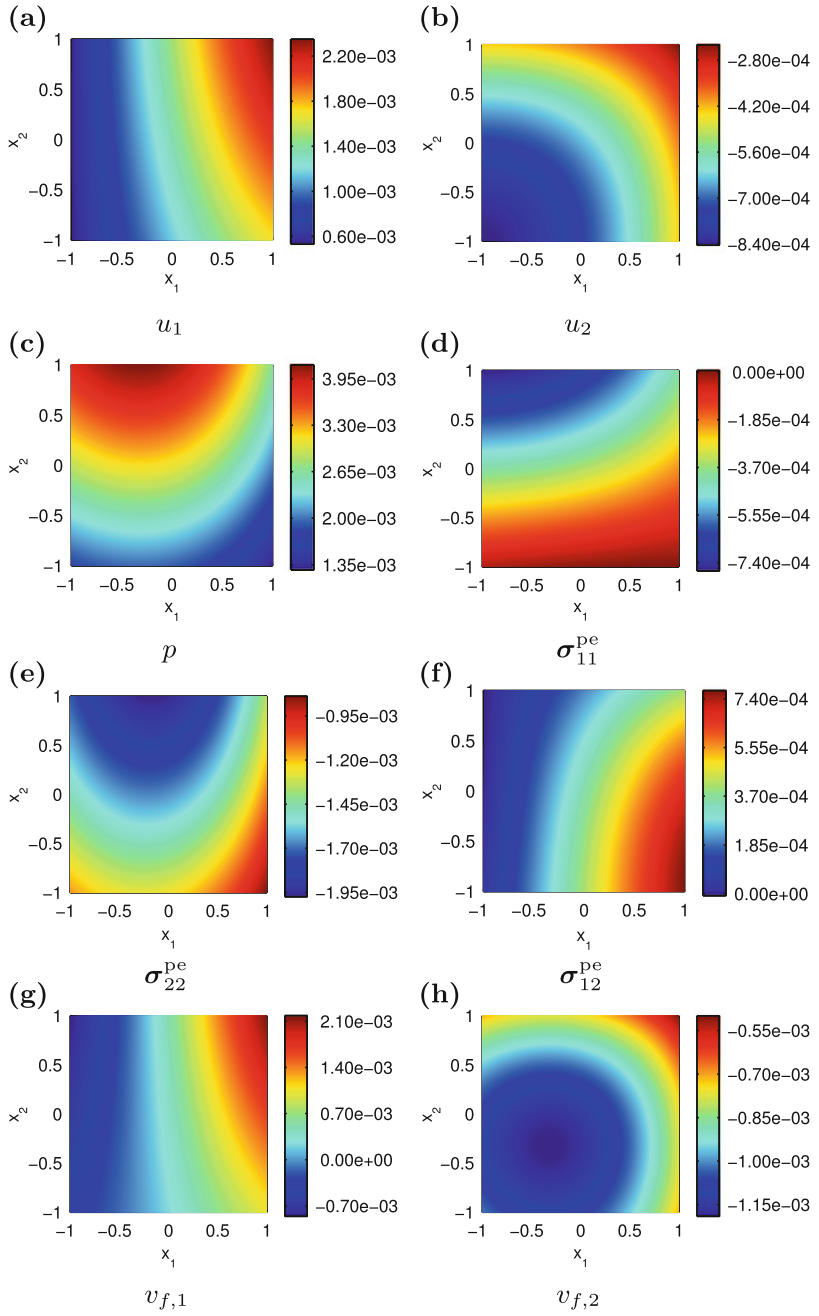

Fig. 6.8 Pseudocolor plots of the poroelastic quantities for Examples 6.4 and 6.5 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Please note that different scales with different signs are used for different quantities as indicated by the different scales on the colorbars. All quantities are dimensionless

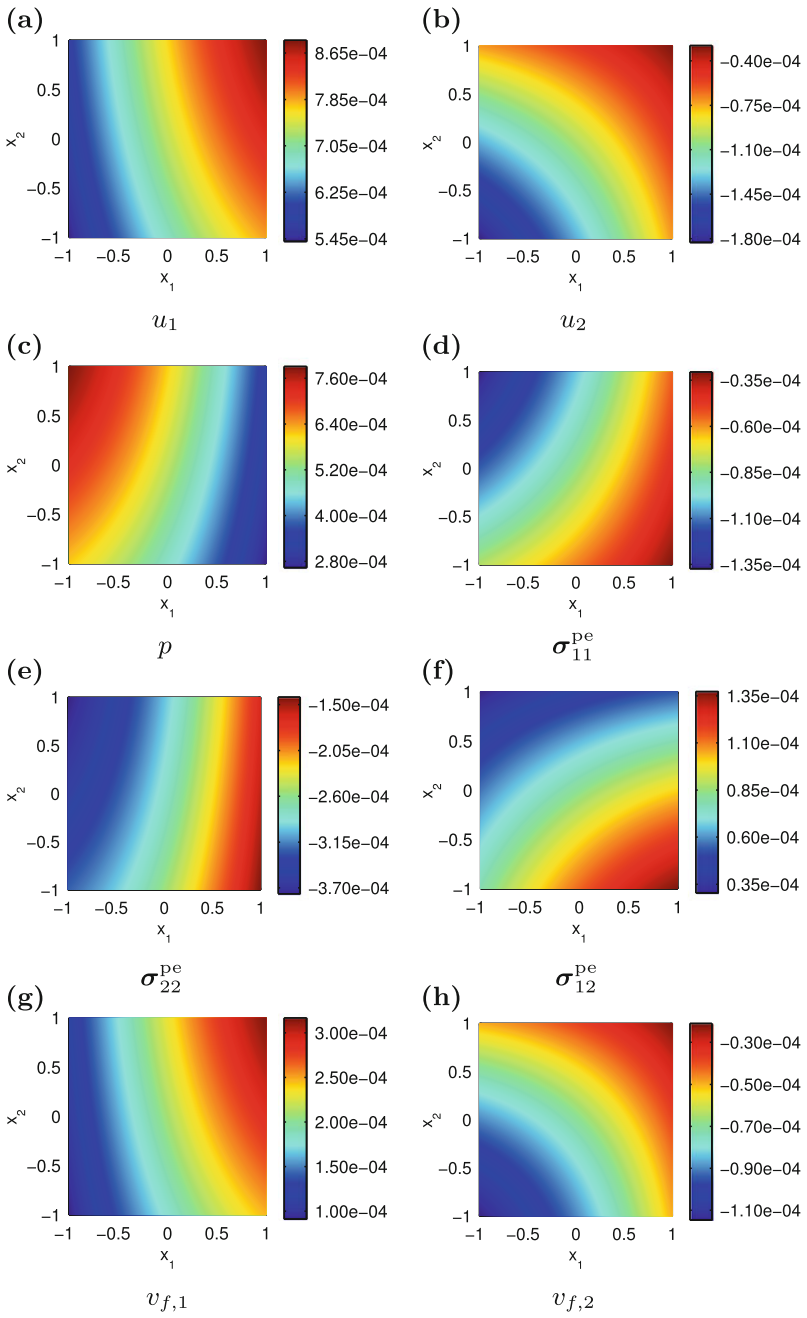

Fig. 6.9 Pseudocolor plots of the poroelastic quantities for Examples 6.4 and 6.5 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Please note that different scales with different signs are used for different quantities as indicated by the different scales on the colorbars. All quantities are dimensionless

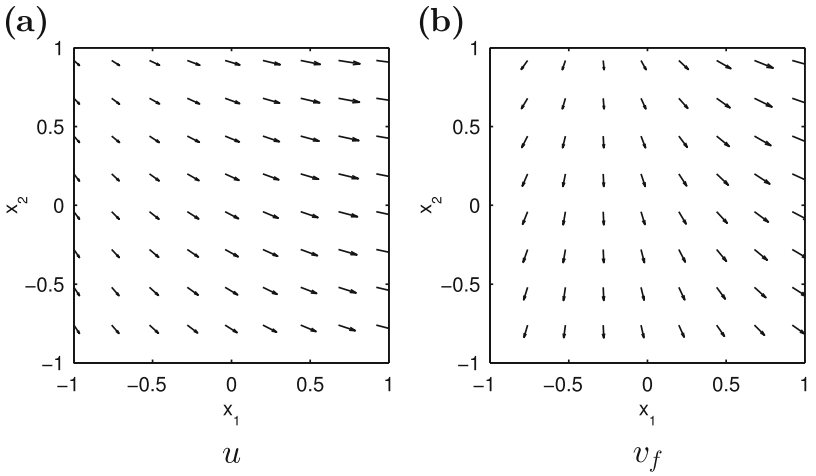

Fig. 6.10 Directions of the vector-valued poroelastic quantities u and v_f for Examples 6.4 and 6.5 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Please compare Fig. 6.8a, b and g, h as the above plots do not give sufficient information about the strength of the fields

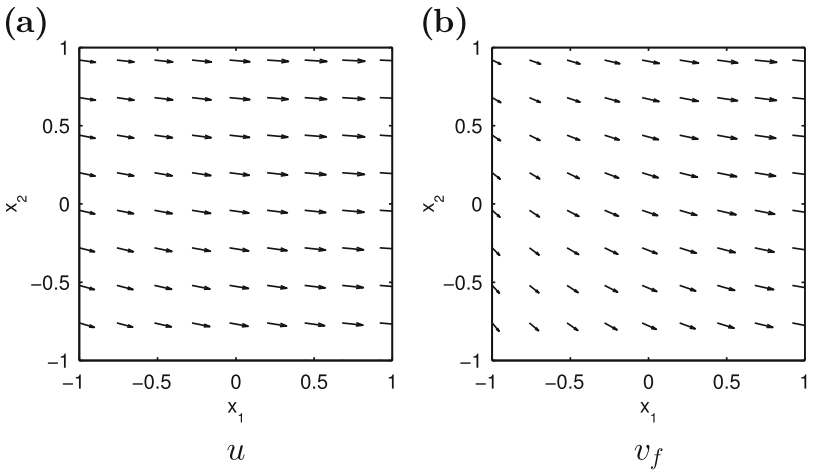

Fig. 6.11 Directions of the vector-valued poroelastic quantities u and v_f for Examples 6.4 and 6.5 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Please compare Fig. 6.9a, b and g, h as the above plots do not give sufficient information about the strength of the fields

Table 6.3 Size of the system matrix as given in (6.7a) for the Δ_x, Δ_y according to (6.14) and $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{\text{end}} = 5$

Δ_x	I	# rows	Δ_y	M	# columns
1	2	3,184	1	2	3,200
1/2	4	6,400	1/2	4	6,400
1/3	6	9,504	1/3	6	9,600
1/4	8	12,736	1/4	8	12,800
1/5	10	15,969	1/5	10	16,000

Numerical solutions for Examples 6.2–6.5 are computed for each combination of Δ_x, Δ_y, and γ taken from the sets

$$\Delta_x\,,\Delta_y \in \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}\right\} \quad , \quad \gamma \in \{0.5, 1, 1.5, 2, 2.5, 3\}\,. \tag{6.14}$$

The set for γ is chosen in accordance with Johansson et al. [143] who noted a loss in accuracy for $\gamma < 0.25$ and $\gamma > 4$ and Alves [8] who found good results with his source point choice for small spatial distances between Γ and $\hat{\Gamma}$.

Table 6.3 gives the number of rows and columns of the system matrix given in (6.7a) for Δ_x, Δ_y according to (6.14). The spectral condition number of the matrix, given by the ratio of its smallest to its largest singular value, ranges between $3.1 \cdot 10^{16}$ and $8.3 \cdot 10^{17}$ for Dirichlet boundary conditions and between $2.3 \cdot 10^{16}$ and $9.1 \cdot 10^{17}$ for mixed boundary conditions.

Figure 6.12 shows $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}$ in case of Example 6.2 for the parameter combination given by values of I and M corresponding to Δ_x, Δ_y as given in (6.14) (see also Table 6.3). The value of γ is color coded by the color of the dots as explained in the caption of the figure. The corresponding relative rooted mean square error is maximized over the points of $\mathfrak{I}_1$ as given in (6.12) and minimized over γ, i.e.

$$\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}} = \min_{\gamma \in \{0.5,1,1.5,2,2.5,3\}} \max_{x \in \mathfrak{I}_1} \mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}(x)\,. \tag{6.15}$$

Figures 6.13–6.15 give $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}$ for Examples 6.3–6.5, respectively. The maximal relative error $\mathrm{E}^{\mathrm{rel}}_{\mathrm{max}}$, evaluated on $\mathfrak{I}_1 \times \{t = j\Delta_t, j \in \mathbb{N}, j \le J\}$ shows the same behavior. Thus, we decided to not depict it here for the sake of plot readability. Moreover, we refrain from denoting at which points of $\mathfrak{I}_1$ and $\mathfrak{I}_t$, respectively, those maximal errors $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}$ are achieved, as error distribution in time and in domain is discussed in more detail in Sect. 6.2.4.

Figure 6.16 shows $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}$ in case of Example 6.2 for the parameter combination given by values of I and M corresponding to Δ_x, Δ_y as given in (6.14). The value of γ is again color coded by the color of the dots as explained in the caption of the figure. The corresponding relative rooted mean square error is maximized over the points of $\mathfrak{I}_t$ as given in (6.11) and minimized over γ, i.e.

$$\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}} = \min_{\gamma \in \{0.5,1,1.5,2,2.5,3\}} \max_{t \in \mathfrak{I}_t} \mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}(t)\,. \tag{6.16}$$

Figures 6.17–6.19 give $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}$ for Examples 6.3–6.5, respectively. As before, the maximal relative error $\mathrm{E}^{\mathrm{rel}}_{\mathrm{max}}$, evaluated on $\mathfrak{I}_2 \times \mathfrak{I}_t$ shows the same behavior and is not depicted.

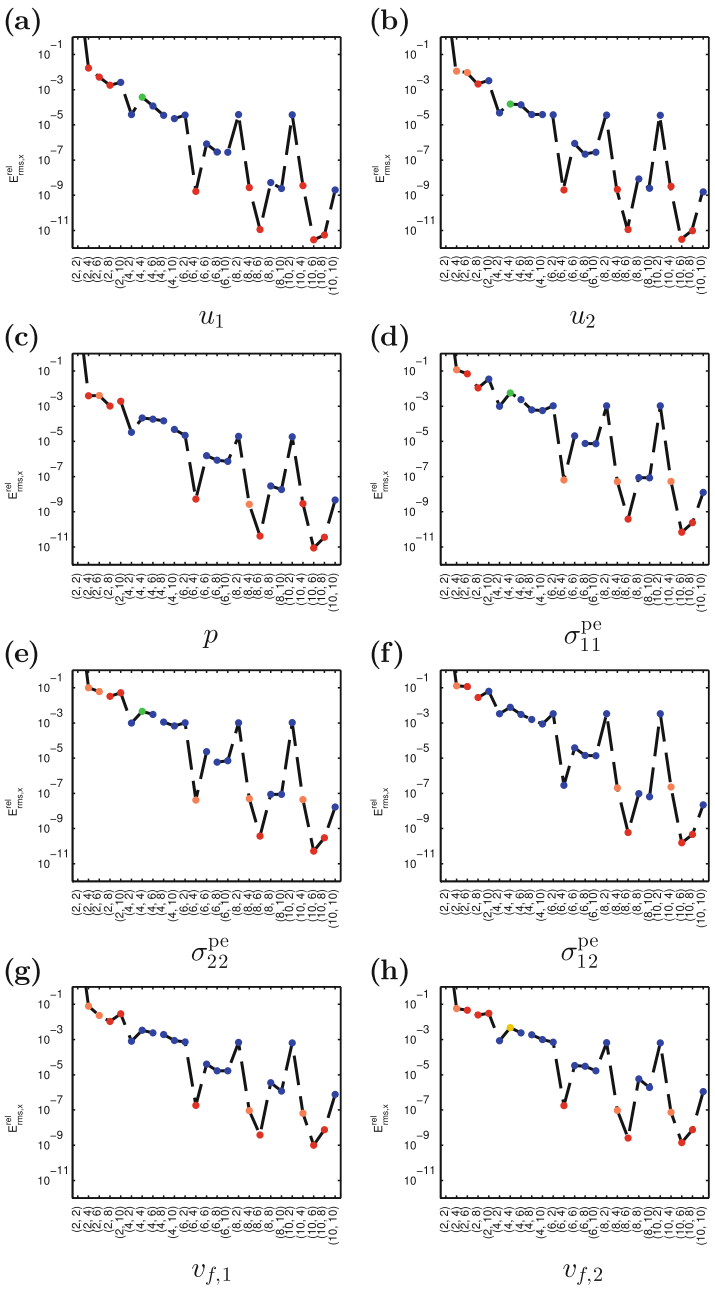

Fig. 6.12 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.2 for varying spatial parameters Δ_x, Δ_y, and γ as given in (6.14) and fixed values $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{end} = 5$. Instead of (Δ_x, Δ_y), we use the corresponding (I, M) (see Table 6.3) on the abscissa. The values of γ are color coded in the following way: • $\gamma = 1.0$, • $\gamma = 1.5$, • $\gamma = 2.0$, • $\gamma = 2.5$, • $\gamma = 3.0$. The best errors were never achieved at $\gamma = 0.5$

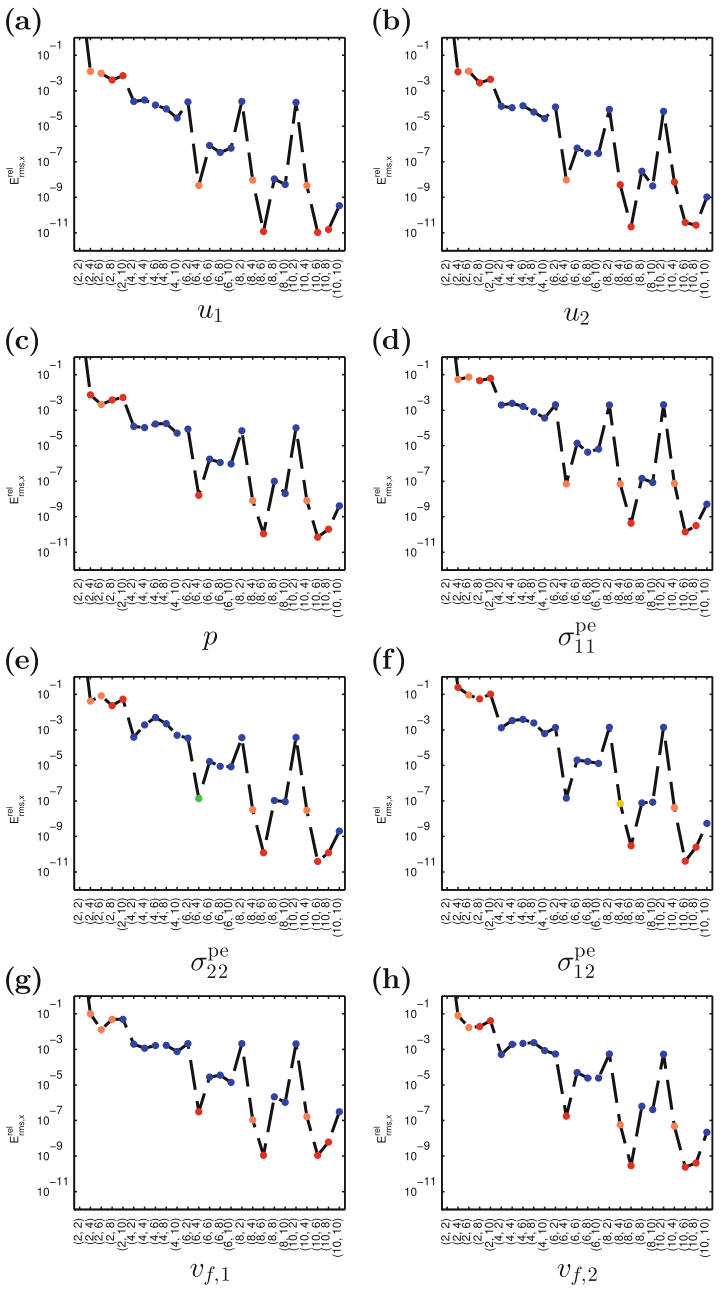

Fig. 6.13 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.3 for varying spatial parameters Δ_x, Δ_y, and γ as given in (6.14) and fixed values $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{end} = 5$. Instead of (Δ_x, Δ_y), we use the corresponding (I, M) (see Table 6.3) on the abscissa. The values of γ are color coded in the following way: • $\gamma = 1.0$, • $\gamma = 1.5$, • $\gamma = 2.0$, • $\gamma = 2.5$, • $\gamma = 3.0$. The best errors were never achieved at $\gamma = 0.5$

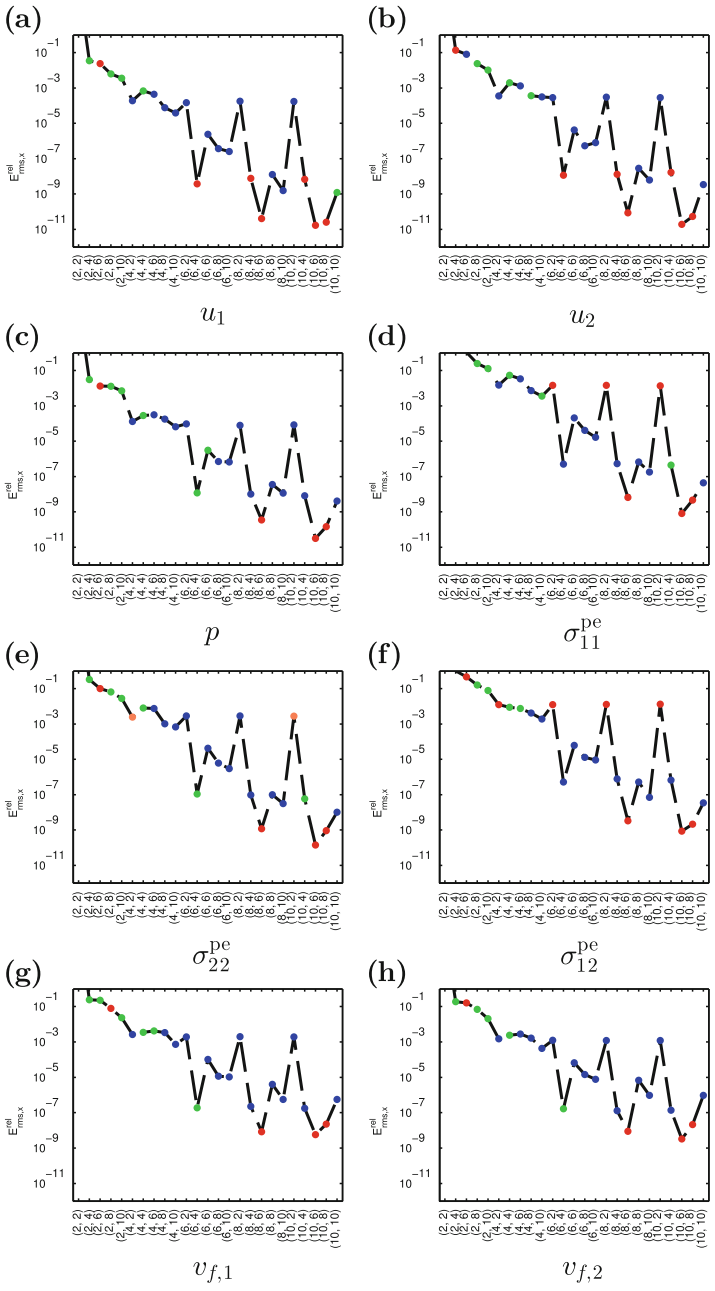

Fig. 6.14 Maximal relative rooted mean square errors $E_{\mathrm{rms,t}}^{\mathrm{rel}}$ as defined in (6.15) for Example 6.4 for varying spatial parameters Δ_x, Δ_y, and γ as given in (6.14) and fixed values $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$. Instead of (Δ_x, Δ_y), we use the corresponding (I, M) (see Table 6.3) on the abscissa. The values of γ are color coded in the following way: • $\gamma = 1.0$, • $\gamma = 1.5$, • $\gamma = 2.0$, • $\gamma = 2.5$, • $\gamma = 3.0$. The best errors were never achieved at $\gamma = 0.5$

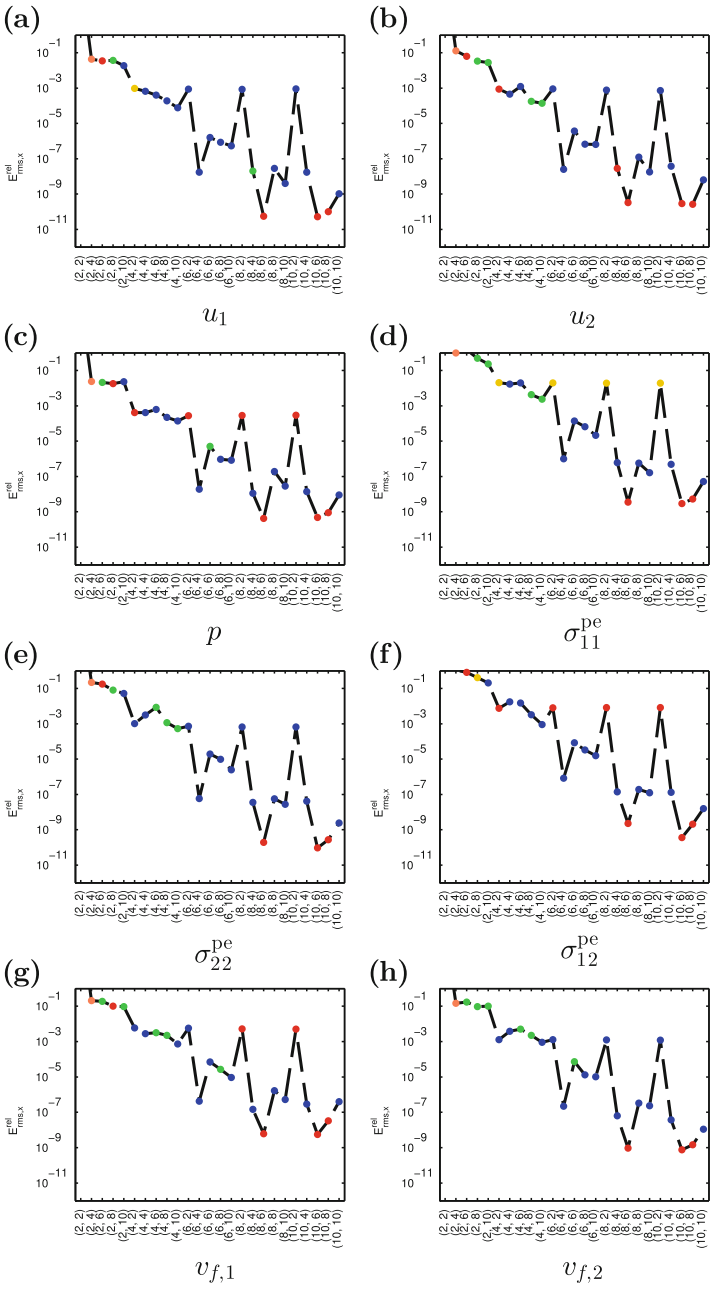

Fig. 6.15 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.5 for varying spatial parameters Δ_x, Δ_y, and γ as given in (6.14) and fixed values $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{end} = 5$. Instead of (Δ_x, Δ_y), we use the corresponding (I, M) (see Table 6.3) on the abscissa. The values of γ are color coded in the following way: • $\gamma = 1.0$, • $\gamma = 1.5$, • $\gamma = 2.0$, • $\gamma = 2.5$, • $\gamma = 3.0$. The best errors were never achieved at $\gamma = 0.5$

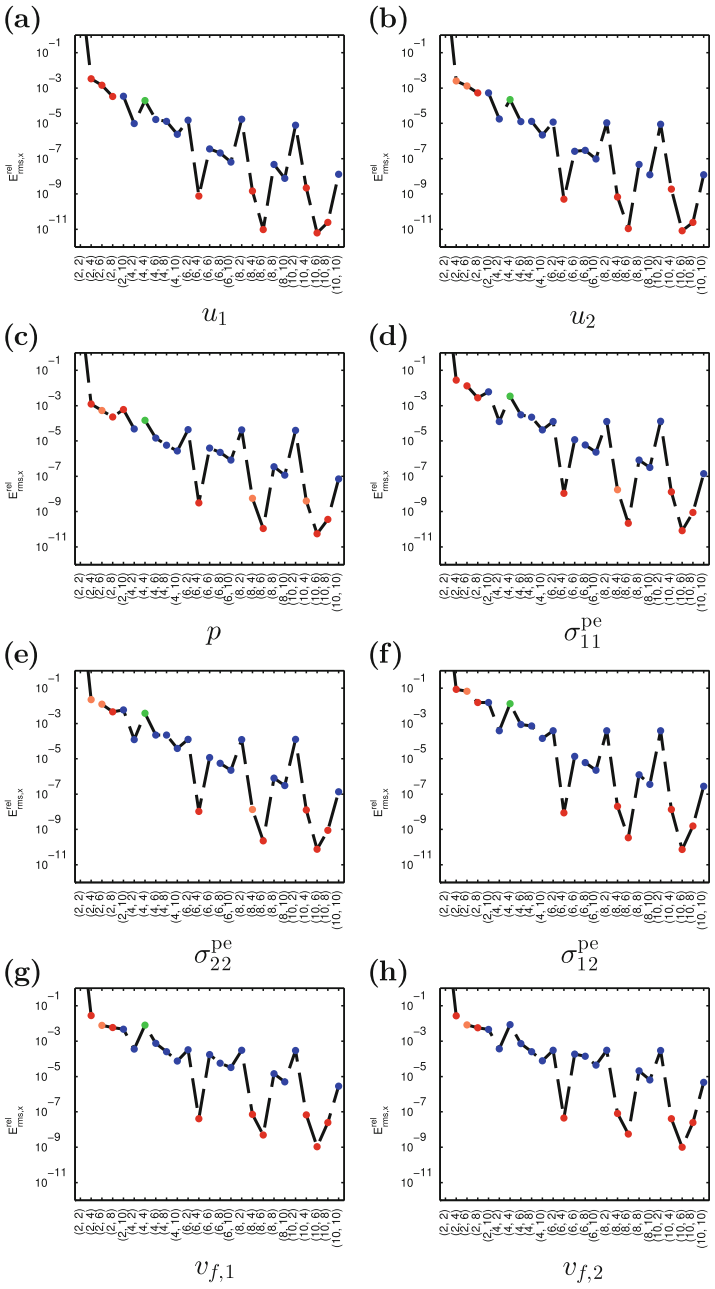

Fig. 6.16 Maximal relative rooted mean square errors $E_{\mathrm{rms,x}}^{\mathrm{rel}}$ as defined in (6.16) for Example 6.2 for varying spatial parameters Δ_x, Δ_y, and γ as given in (6.14) and fixed values $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$. Instead of (Δ_x, Δ_y), we use the corresponding (I, M) (see Table 6.3) on the abscissa. The values of γ are color coded in the following way: • $\gamma = 1.0$, • $\gamma = 1.5$, • $\gamma = 2.0$, • $\gamma = 2.5$, • $\gamma = 3.0$. The best errors were never achieved at $\gamma = 0.5$

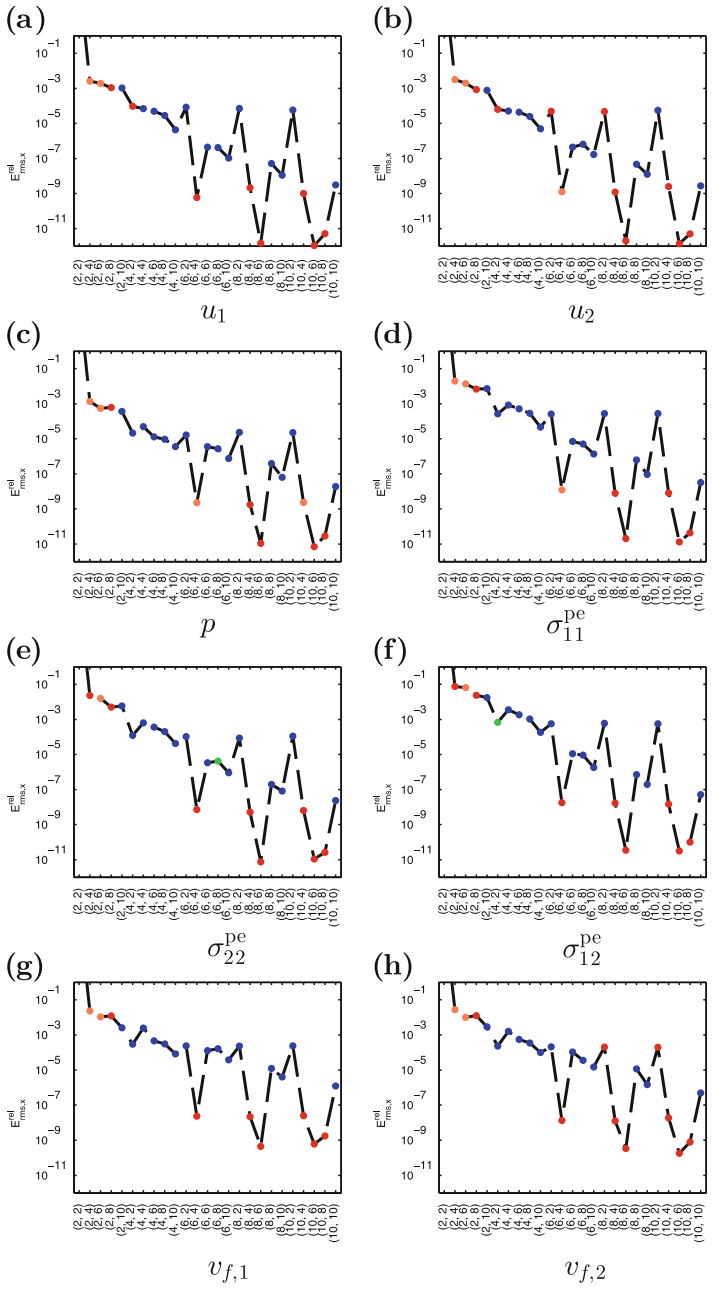

Fig. 6.17 Maximal relative rooted mean square errors $E_{\mathrm{rms,x}}^{\mathrm{rel}}$ as defined in (6.16) for Example 6.3 for varying spatial parameters Δ_x, Δ_y, and γ as given in (6.14) and fixed values $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$. Instead of (Δ_x, Δ_y), we use the corresponding (I, M) (see Table 6.3) on the abscissa. The values of γ are color coded in the following way: • $\gamma = 1.0$, • $\gamma = 1.5$, • $\gamma = 2.0$, • $\gamma = 2.5$, • $\gamma = 3.0$. The best errors were never achieved at $\gamma = 0.5$

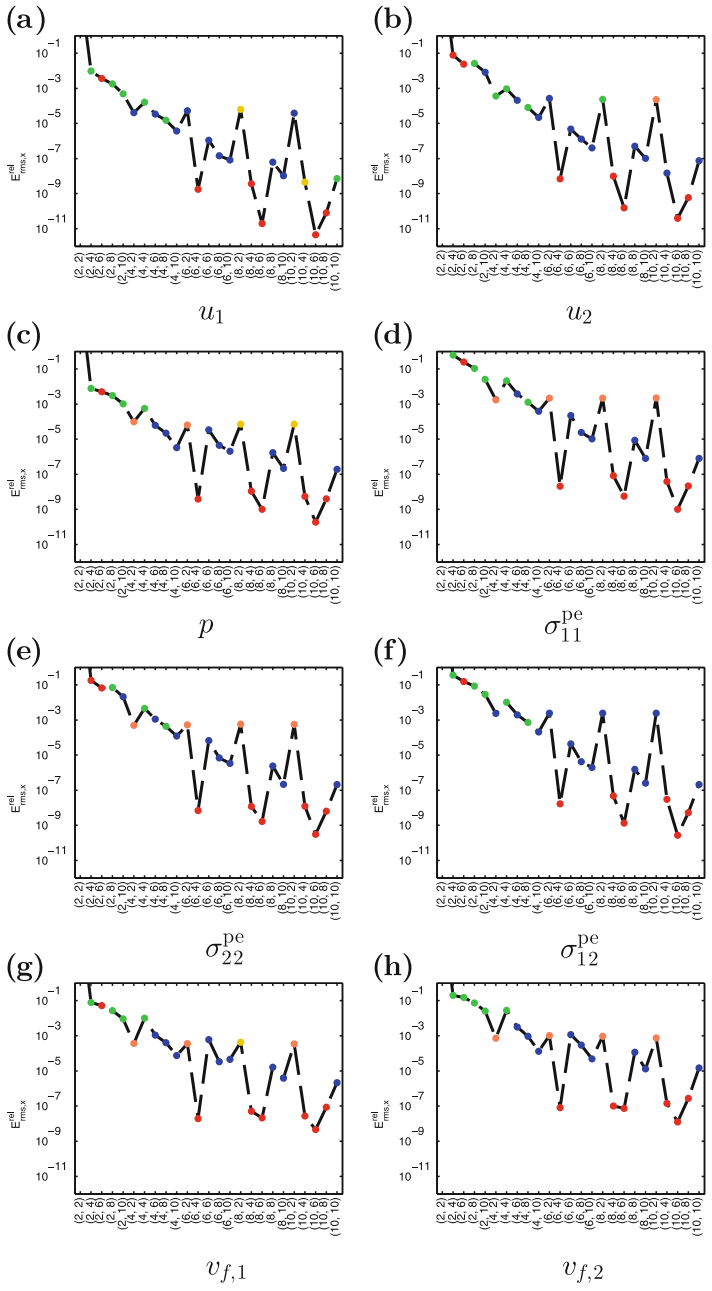

Fig. 6.18 Maximal relative rooted mean square errors $E_{\mathrm{rms,x}}^{\mathrm{rel}}$ as defined in (6.16) for Example 6.4 for varying spatial parameters Δ_x, Δ_y, and γ as given in (6.14) and fixed values $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$. Instead of (Δ_x, Δ_y), we use the corresponding (I, M) (see Table 6.3) on the abscissa. The values of γ are color coded in the following way: • $\gamma = 1.0$, • $\gamma = 1.5$, • $\gamma = 2.0$, • $\gamma = 2.5$, • $\gamma = 3.0$. The best errors were never achieved at $\gamma = 0.5$

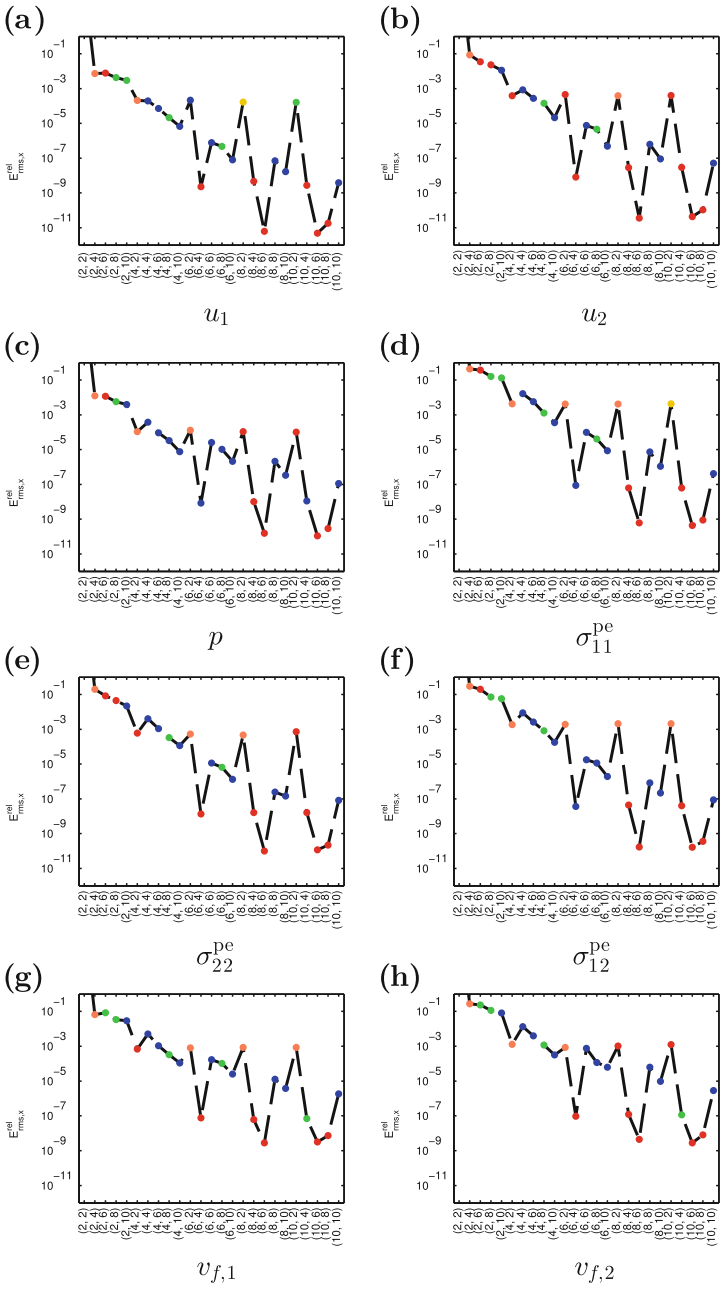

Fig. 6.19 Maximal relative rooted mean square errors $E_{\mathrm{rms,x}}^{\mathrm{rel}}$ as defined in (6.16) for Example 6.5 for varying spatial parameters Δ_x, Δ_y, and γ as given in (6.14) and fixed values $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$. Instead of (Δ_x, Δ_y), we use the corresponding (I, M) (see Table 6.3) on the abscissa. The values of γ are color coded in the following way: • $\gamma = 1.0$, • $\gamma = 1.5$, • $\gamma = 2.0$, • $\gamma = 2.5$, • $\gamma = 3.0$. The best errors were never achieved at $\gamma = 0.5$

The same logarithmic scale is used in all subfigures of Figs. 6.12–6.19 to make all errors easily comparable.

As the qualitative behavior of the error is the same for all subfigures of Figs. 6.12–6.19, i.e., for all quantities in all examples under consideration, we do not discuss them separately. In general, errors in case of Examples 6.3 and 6.5 with mixed boundary conditions are larger than errors for Examples 6.2 and 6.4 with pure Dirichlet boundary conditions. Likewise, errors for the approximation of fi-parts of fundamental solutions (Examples 6.4 and 6.5) are larger than errors for the approximation of Si-parts (Examples 6.2 and 6.3). This becomes more obvious in Sect. 6.2.4.

As can be expected, there is a trend that errors reduce as the number of collocation points (corresponding to I) or the number of source points (corresponding to M) increases. However, there are some remarkable anomalies for parameter combinations $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$ ($I = 6$, $M = 4$), $\Delta_x = \frac{1}{4}$, $\Delta_y = \frac{1}{3}$ ($I = 8$, $M = 6$), $\Delta_x = \frac{1}{5}$, $\Delta_y = \frac{1}{3}$ ($I = 10$, $M = 6$), and $\Delta_x = \frac{1}{5}$, $\Delta_y = \frac{1}{4}$ ($I = 10$, $M = 8$). Here, we find local minima of the relative errors. Additionally, the values of γ which minimize the error for those combinations of Δ_x and Δ_y are unexpectedly small ($\gamma = 1$ in most cases), even if only a small number of source and collocation points is used. This is in contrast to other pairs of Δ_x and Δ_y, which achieve the smallest errors for larger γ ($\gamma = 3$ in most cases). The usual expectation is, that, for small numbers of source points, a larger distance between the domain boundary and the pseudo-boundary gives better results due to better approximation properties, whereas larger sets of source points need to be placed closer to the boundary to counteract the ill-conditioning of the system with growing number of source points and growing distance between Γ and $\hat{\Gamma}$. Moreover, we find the best approximations for overdetermined systems, in contrast to [266, 267] in which good results for underdetermined systems are reported.

In conclusion, the results of this section clearly demonstrate that a suitable choice of Δ_x, Δ_y, and γ can yield remarkably good results, whereas a poor choice leads to errors which are several orders of magnitude larger than the optimum.

6.2.2 *Varying Temporal Parameters for Negative Times*

Next, we investigate the influence of the parameters τ_{-1} and Δ_τ which are relevant for the placement in time of source points for Si-parts of fundamental solutions stemming from the incorporation of non-vanishing boundary conditions. All other parameters are fixed as $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$ in this section, yielding $I = 6$, $M = 4$, and $J = 100$. Therefore, the number of collocation and source points do not change within this section. Only the placement of those source points with negative τ is varied.

We have seen in the previous section that the spatial parameters chosen here yield a remarkably good approximation with magnitudes of relative error between 10^{-6} and 10^{-9}, depending on the quantity and example under consideration.

Numerical solutions for Examples 6.2–6.5 are computed for each combination of

$$\tau_{-1}, \Delta_\tau \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1\}. \tag{6.17}$$

As already mentioned, it is shown in [131] that choosing all source points with negative τ close to $t = 0$ results in large approximation errors. A possible explanation for this could be that making the time interval in which such source points may be chosen too small leads to matrices with many very similar column vectors, thus creating instabilities in a numerical solution scheme. This would imply that the main reason for large approximation errors is not primarily the placement of the source points in time but the temporal distance between them.

Figure 6.20 shows $E^{\text{rel}}_{\text{rms,t}}$ in case of Example 6.2. The abscissa is partitioned into seven segments according to the seven values of τ_{-1} under consideration. In each segment, Δ_τ varies from small to large from left to right. As γ is not varied, there is no minimization in Eqs. (6.15) and (6.16) which define the maximum rooted mean square errors. Figures 6.21–6.23 give $E^{\text{rel}}_{\text{rms,t}}$ for Examples 6.3–6.5, respectively, for the considered combinations of τ_{-1} and Δ_τ. The maximum rooted mean square error $E^{\text{rel}}_{\text{rms,x}}$ in the domain is shown in Figs. 6.24–6.27.

Figures 6.20–6.27 use the same logarithmic scale as Figs. 6.12–6.19 for all poroelastic quantities in order to simplify comparisons of the influence of different kinds of parameters.

It is evident from Figs. 6.20–6.27 that the influence of τ_{-1} on the errors is not very strong. Results get slightly worse for large values of τ_{-1}.

In contrast, choosing Δ_τ too small or too large yields significantly larger errors. This is in line with our possible explanation for the results in [131]. Small Δ_τ means that the temporal distance between source points is small, resulting in very similar columns in the system matrix. On the other hand, in case of very large Δ_τ, many source points are placed far away in time from $\Omega \times (0, t_{\text{end}})$ which also results in many similar columns. Especially in Figs. 6.24–6.27, showing the error in the domain, it can be seen that large Δ_τ yield a significant increase in errors of up to three orders of magnitude. The best results are gained for $\Delta_\tau = 0.05$, $\Delta_\tau = 0.1$, and $\Delta_\tau = 0.2$. Which of these three values is actually chosen has only a minor influence on the overall result.

In conclusion of this section, we find that unreasonable choices of Δ_τ, i.e., rather small or rather large values, increase the approximation errors, whereas if Δ_τ is chosen within appropriate limits, the actual value is not that important. Moreover, choosing τ_{-1} is even less critical. In comparison, choosing suitable spatial method parameters (Δ_x, Δ_y, and γ) is much more critical and less intuitive.

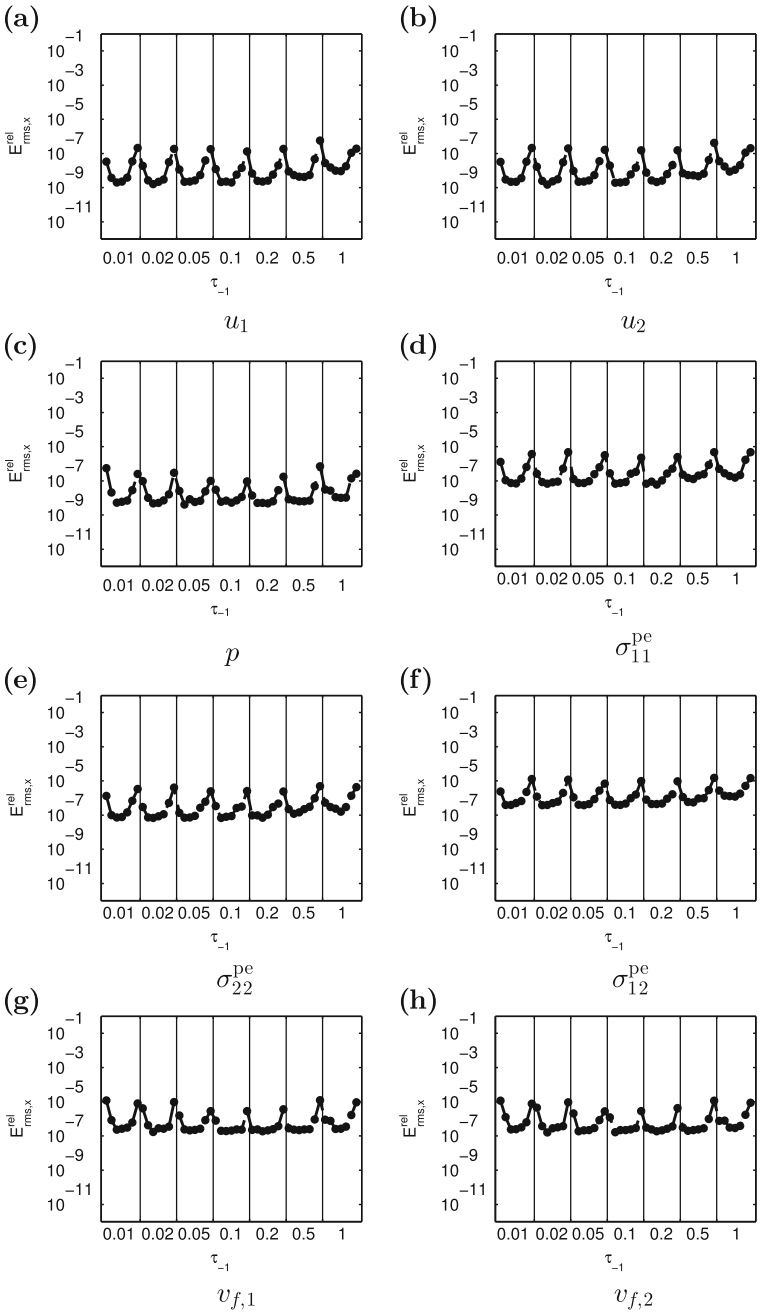

Fig. 6.20 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.2 for varying parameters τ_{-1}, Δ_τ as given in (6.17) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{end} = 5$. Each plot is split into seven parts with the corresponding τ_{-1} given on the abscissa. For each part, Δ_τ varies from small to large from left to right

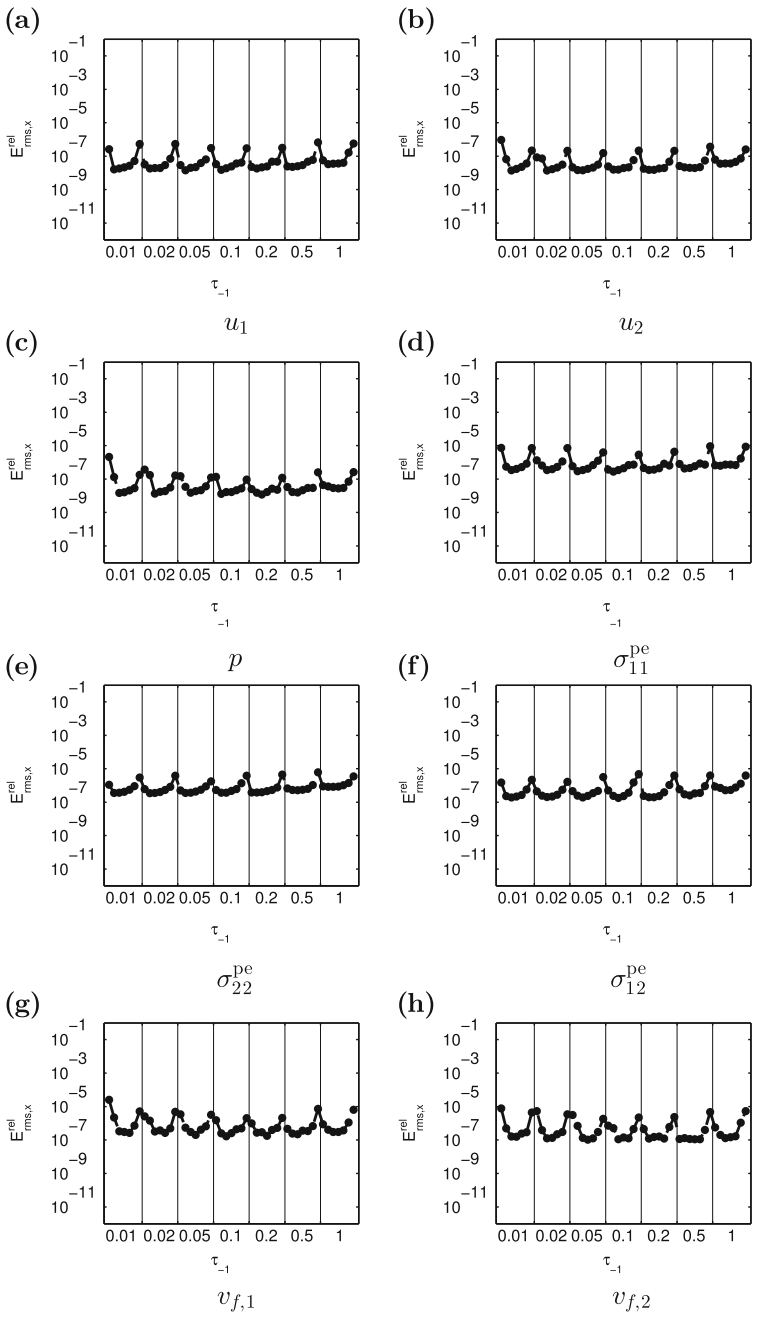

Fig. 6.21 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.3 for varying parameters τ_{-1}, Δ_τ as given in (6.17) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{end} = 5$. Each plot is split into seven parts with the corresponding τ_{-1} given on the abscissa. For each part, Δ_τ varies from small to large from left to right

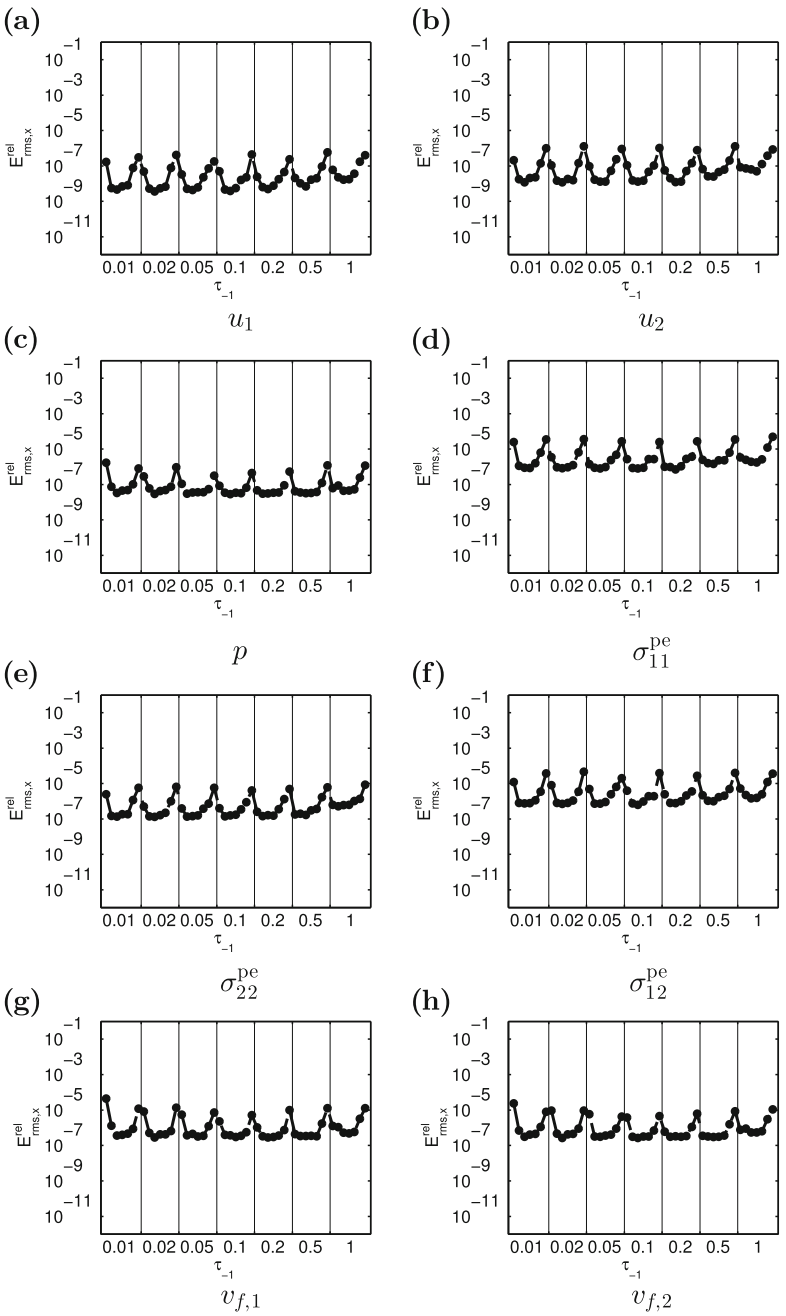

Fig. 6.22 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.4 for varying parameters τ_{-1}, Δ_τ as given in (6.17) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{end} = 5$. Each plot is split into seven parts with the corresponding τ_{-1} given on the abscissa. For each part, Δ_τ varies from small to large from left to right

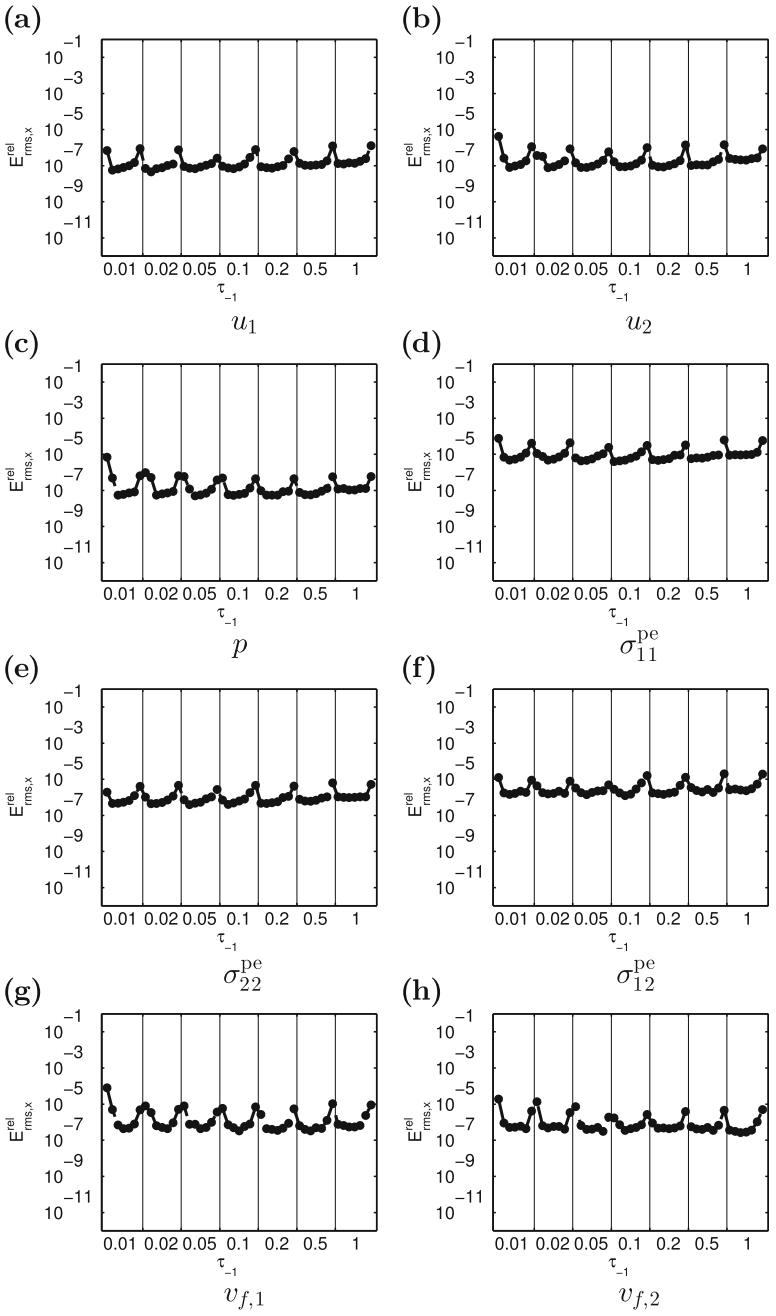

Fig. 6.23 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.5 for varying parameters τ_{-1}, Δ_τ as given in (6.17) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{end} = 5$. Each plot is split into seven parts with the corresponding τ_{-1} given on the abscissa. For each part, Δ_τ varies from small to large from left to right

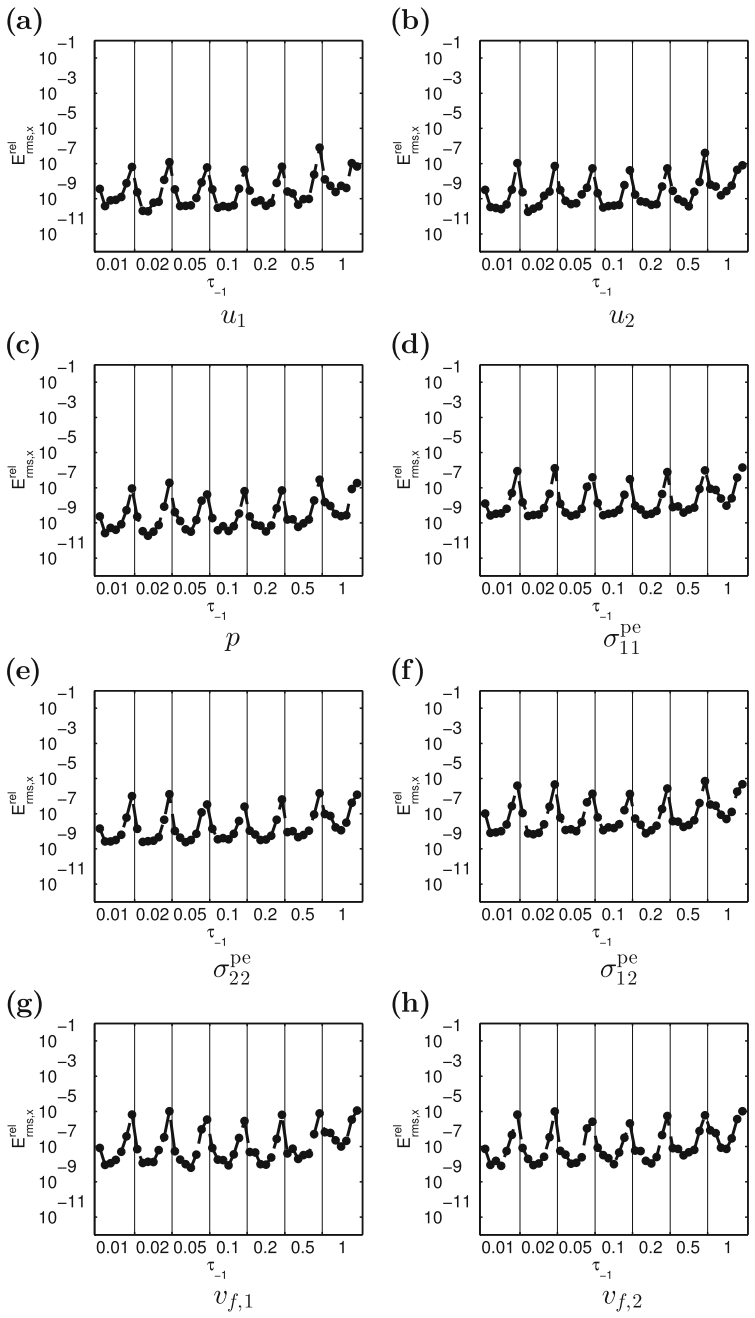

Fig. 6.24 Maximal relative rooted mean square errors $E^{rel}_{rms,x}$ as defined in (6.16) for Example 6.2 for varying parameters τ_{-1}, Δ_τ as given in (6.17) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{end} = 5$. Each plot is split into seven parts with the corresponding τ_{-1} given on the abscissa. For each part, Δ_τ varies from small to large from left to right

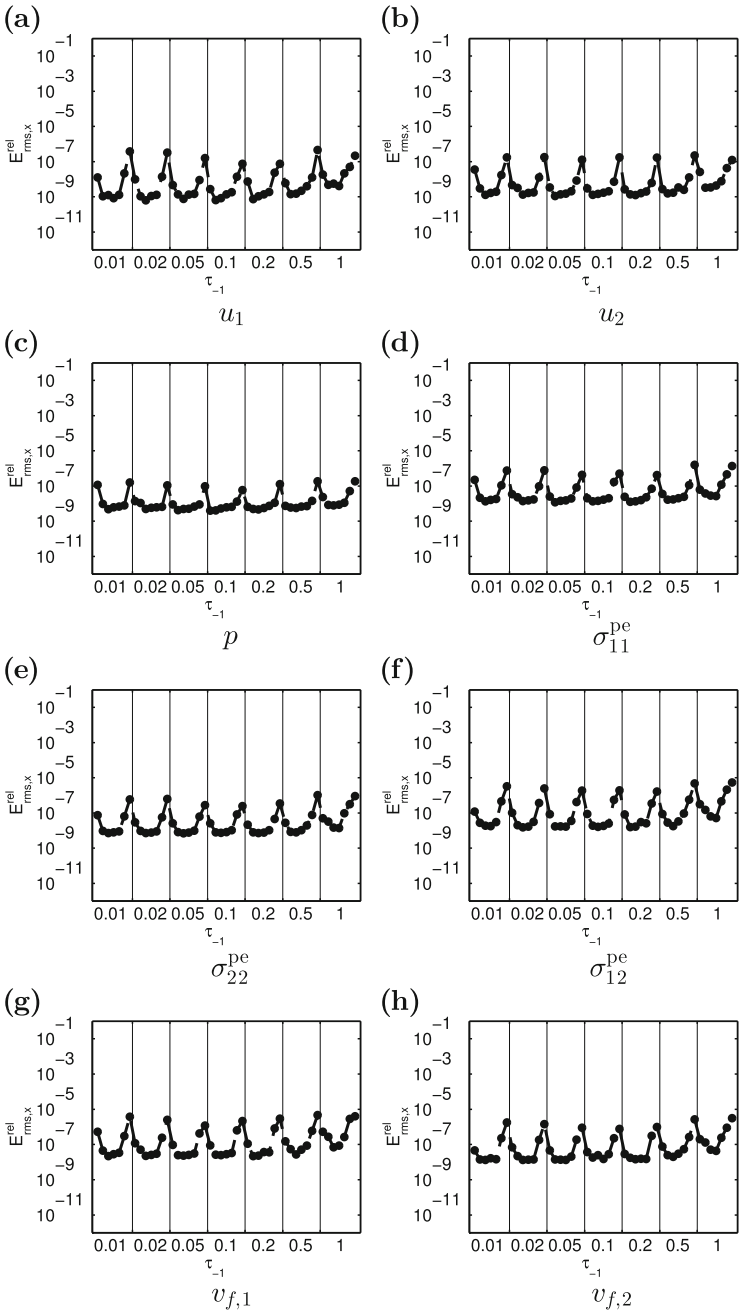

Fig. 6.25 Maximal relative rooted mean square errors $E^{rel}_{rms,x}$ as defined in (6.16) for Example 6.3 for varying parameters τ_{-1}, Δ_τ as given in (6.17) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{end} = 5$. Each plot is split into seven parts with the corresponding τ_{-1} given on the abscissa. For each part, Δ_τ varies from small to large from left to right

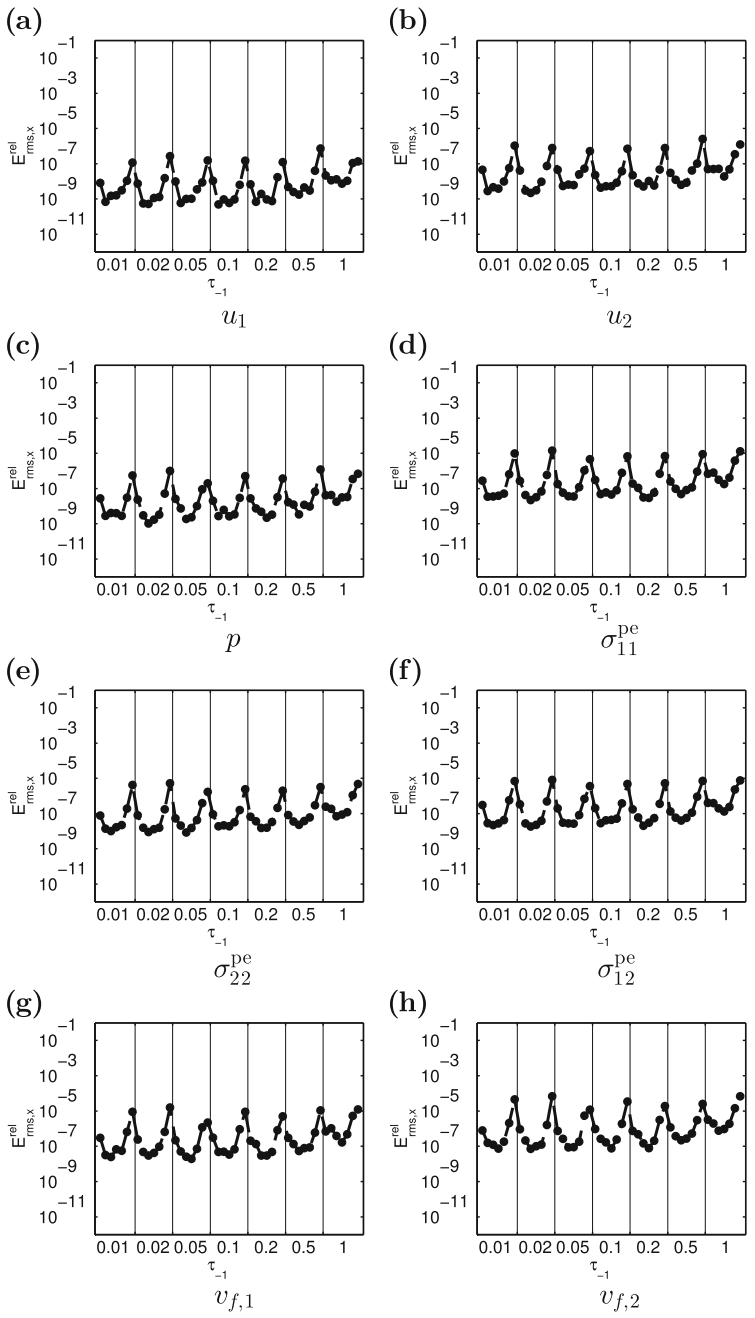

Fig. 6.26 Maximal relative rooted mean square errors $E^{rel}_{rms,x}$ as defined in (6.16) for Example 6.4 for varying parameters τ_{-1}, Δ_τ as given in (6.17) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{end} = 5$. Each plot is split into seven parts with the corresponding τ_{-1} given on the abscissa. For each part, Δ_τ varies from small to large from left to right

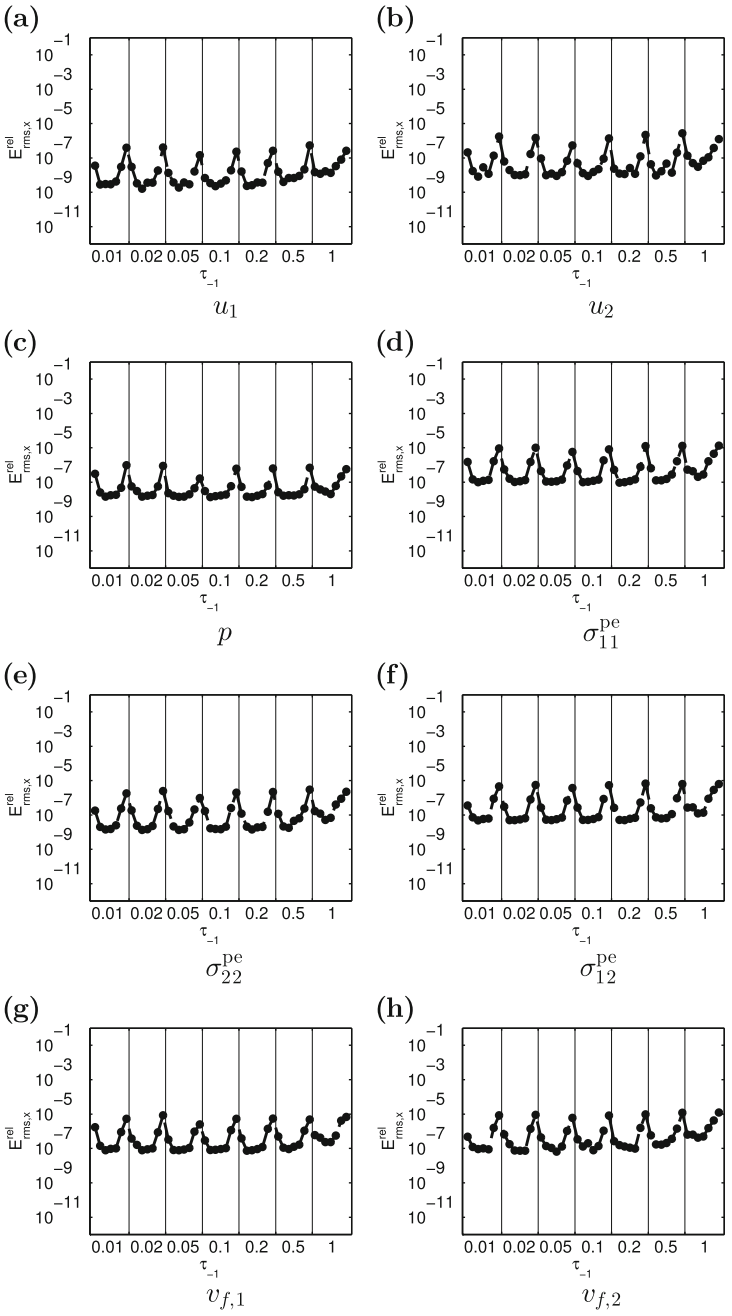

Fig. 6.27 Maximal relative rooted mean square errors $E^{rel}_{rms,x}$ as defined in (6.16) for Example 6.5 for varying parameters τ_{-1}, Δ_τ as given in (6.17) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_t = 0.05$, $\tau_1 = 0.5\Delta_t$, $t_{end} = 5$. Each plot is split into seven parts with the corresponding τ_{-1} given on the abscissa. For each part, Δ_τ varies from small to large from left to right

6.2.3 *Varying Temporal Parameters for Positive Times*

Within this section, we vary the temporal positions of collocation points and source points with positive τ, as we consider different combinations of Δ_t and τ_1. In contrast to the previous sections, we do not vary these two parameters independently, as our method demands $0 < \tau_1 < \Delta_t$. Instead, we test different ratios of τ_1 to Δ_t. The different combinations of Δ_t and τ_1 are given by

$$\begin{aligned} \Delta_t &\in \{0.01, 0.02, 0.05, 0.1, 0.2\}\,, \\ \frac{\tau_1}{\Delta_t} &\in \{0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.99\}\,. \end{aligned} \tag{6.18}$$

Spatial parameters are fixed as $\Delta_x = \frac{1}{3}$ ($I = 6$), $\Delta_y = \frac{1}{2}$ ($M = 4$), and $\gamma = 1$, whereas the remaining temporal parameters are chosen with a fixed ratio to Δ_t as $\tau_{-1} = 0.5\Delta_t$ and $\Delta_\tau = \Delta_t$.

As Δ_t varies, the same is true for J. Thus, we choose a smaller time interval by setting $t_{\text{end}} = 1$ instead of $t_{\text{end}} = 5$ as in the previous examples, as this would yield very large system matrices for $\Delta_t = 0.01$. $\mathfrak{J}_t$ is adapted accordingly to

$$\mathfrak{J}_t^{(1)} = \{0.2, 0.4, 0.6, 0.8, 1.0\}. \tag{6.19}$$

Figure 6.28 shows $\mathrm{E}_{\mathrm{rms,t}}^{\mathrm{rel}}$ in case of Example 6.2. Instead of partitioning the abscissa according to the choice of τ_1, it is preferable here to make a partition according to the five values of Δ_t under consideration, as the choice of τ_1 depends on the choice of Δ_t. In each segment, τ_1 is varied from small to large from left to right according to (6.18). As in the last section, γ is not varied. Therefore, there is no minimization in Eqs. (6.15) and (6.16). Figures 6.29–6.31 give $\mathrm{E}_{\mathrm{rms,t}}^{\mathrm{rel}}$ for Examples 6.3–6.5, respectively, for the considered combinations of Δ_t and τ_1. The maximum error $\mathrm{E}_{\mathrm{rms,x}}^{\mathrm{rel}}$ in the domain is shown in Figs. 6.32–6.35.

As in all previous considerations, Figs. 6.28–6.35 use the same logarithmic scale as Figs. 6.12–6.27 for all poroelastic quantities in order to simplify comparisons of the influence of different kinds of parameters.

As expected, choosing larger values of Δ_t, which results in fewer collocation and source points, increases the error. On the other hand, choosing Δ_t smaller than 0.05 does not necessarily yield smaller errors. The errors' order of magnitude for smaller Δ_t seems to be constant for some quantities (u and σ as seen in subfigures a, b, and d–f of Figs. 6.29 and 6.31), whereas others may even increase (v as seen in subfigures g–h of Figs. 6.29 and 6.31). The influence of the ratio of τ_1 to Δ_t is rather small, with a slight decrease for $\frac{\tau_1}{\Delta_t} = 0.5$ and $\frac{\tau_1}{\Delta_t} = 0.75$ in some cases (e.g., in the pressure p in Figs. 6.32g and 6.31c or the fluid velocity v_f in Figs. 6.30g and 6.32g).

Overall, we conclude that the actual choice of temporal parameters for positive times is the least critical of all three sets under consideration, as long as we bear in mind that choosing a large Δ_t results in few collocation and source points, thus naturally yielding larger approximation errors.

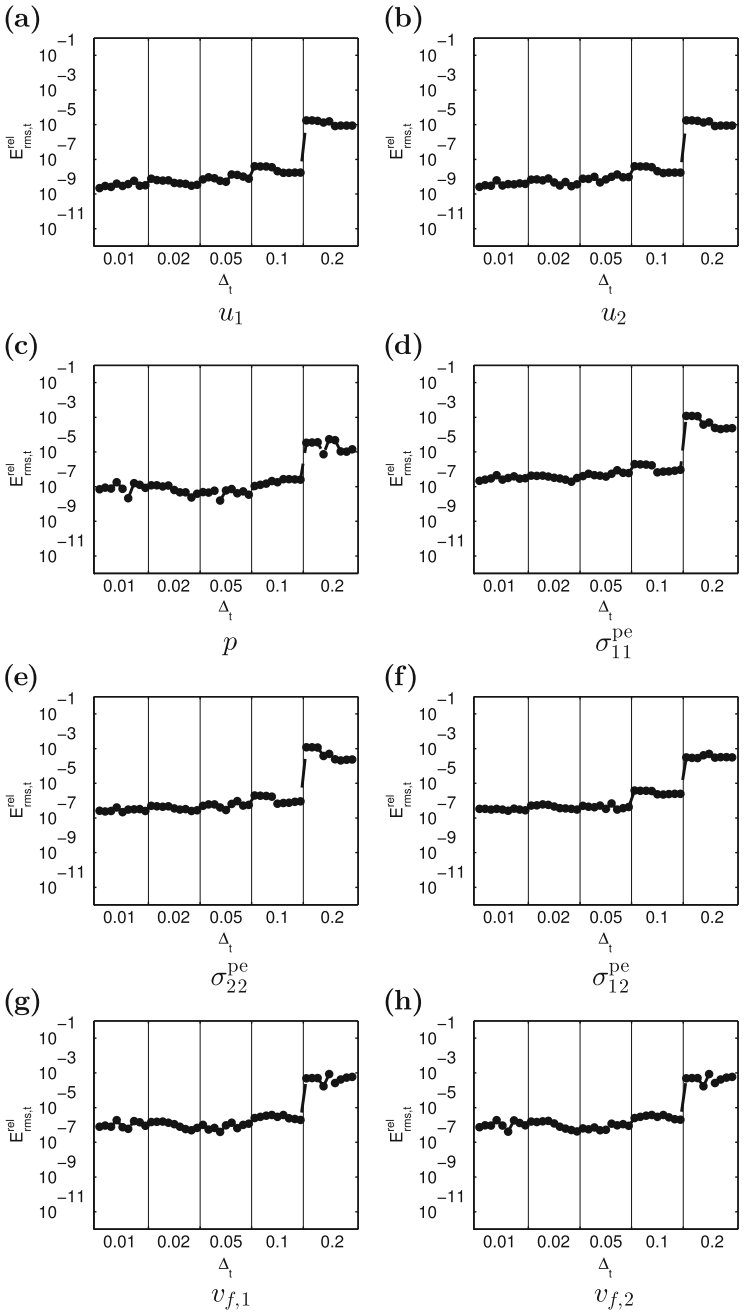

Fig. 6.28 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.2 for varying parameters Δ_t, τ_1 as given in (6.18) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t$, $\tau_{-1} = 0.5\Delta_t$, $t_{end} = 1$. Each plot is split into five parts with the corresponding Δ_t given on the abscissa. For each part, τ_1 varies from small to large from left to right

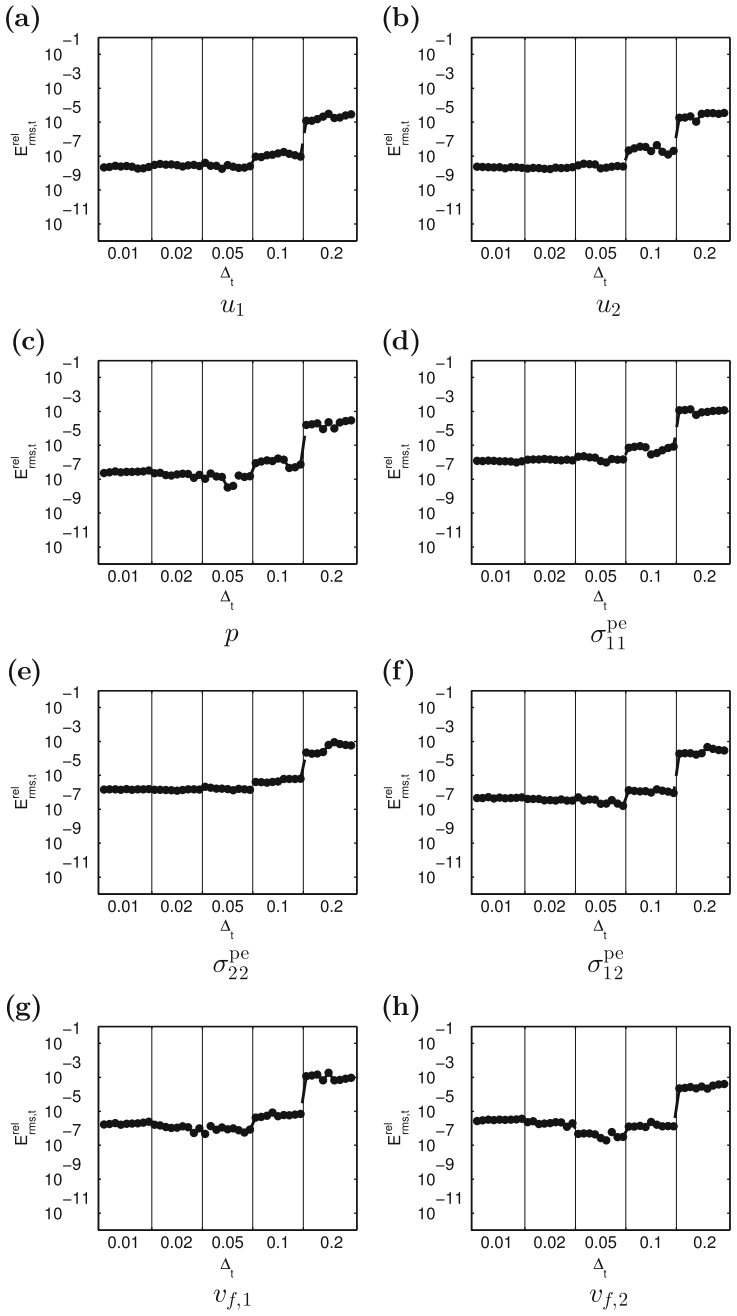

Fig. 6.29 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.3 for varying parameters Δ_t, τ_1 as given in (6.18) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t$, $\tau_{-1} = 0.5\Delta_t$, $t_{end} = 1$. Each plot is split into five parts with the corresponding Δ_t given on the abscissa. For each part, τ_1 varies from small to large from left to right

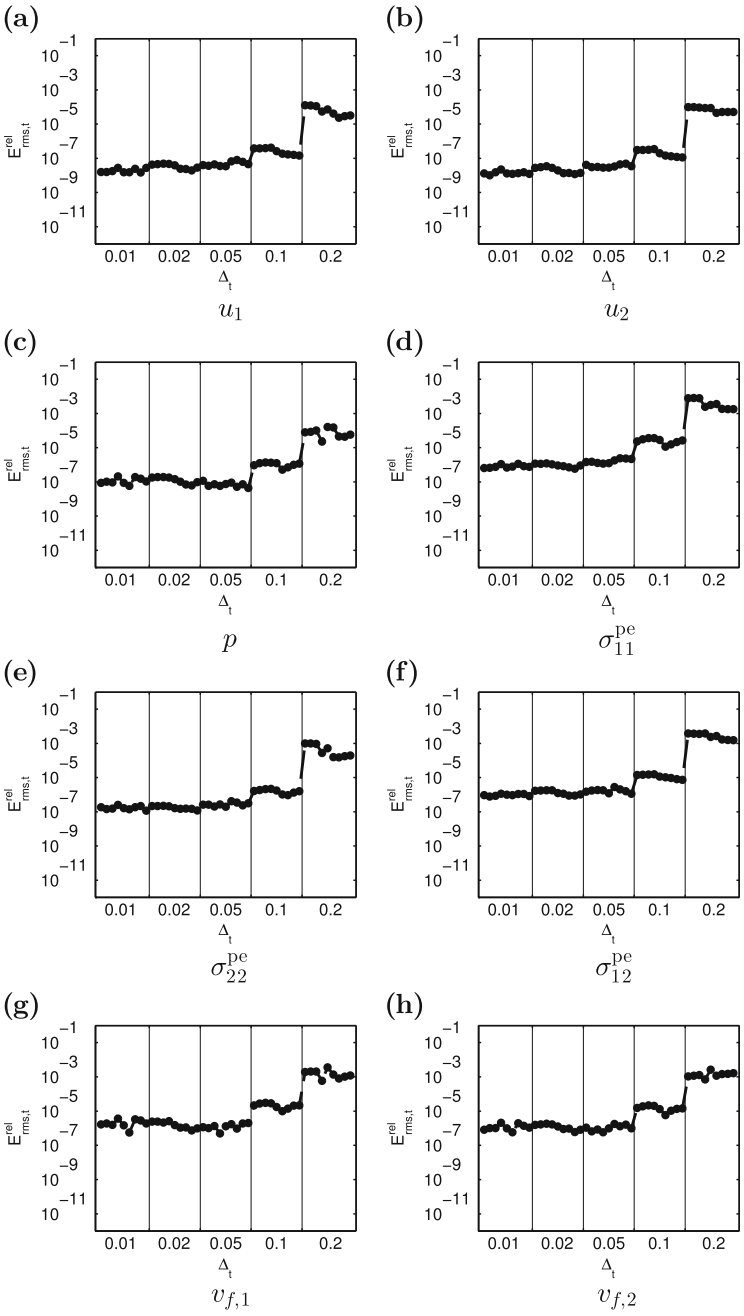

Fig. 6.30 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.4 for varying parameters Δ_t, τ_1 as given in (6.18) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t$, $\tau_{-1} = 0.5\Delta_t$, $t_{end} = 1$. Each plot is split into five parts with the corresponding Δ_t given on the abscissa. For each part, τ_1 varies from small to large from left to right

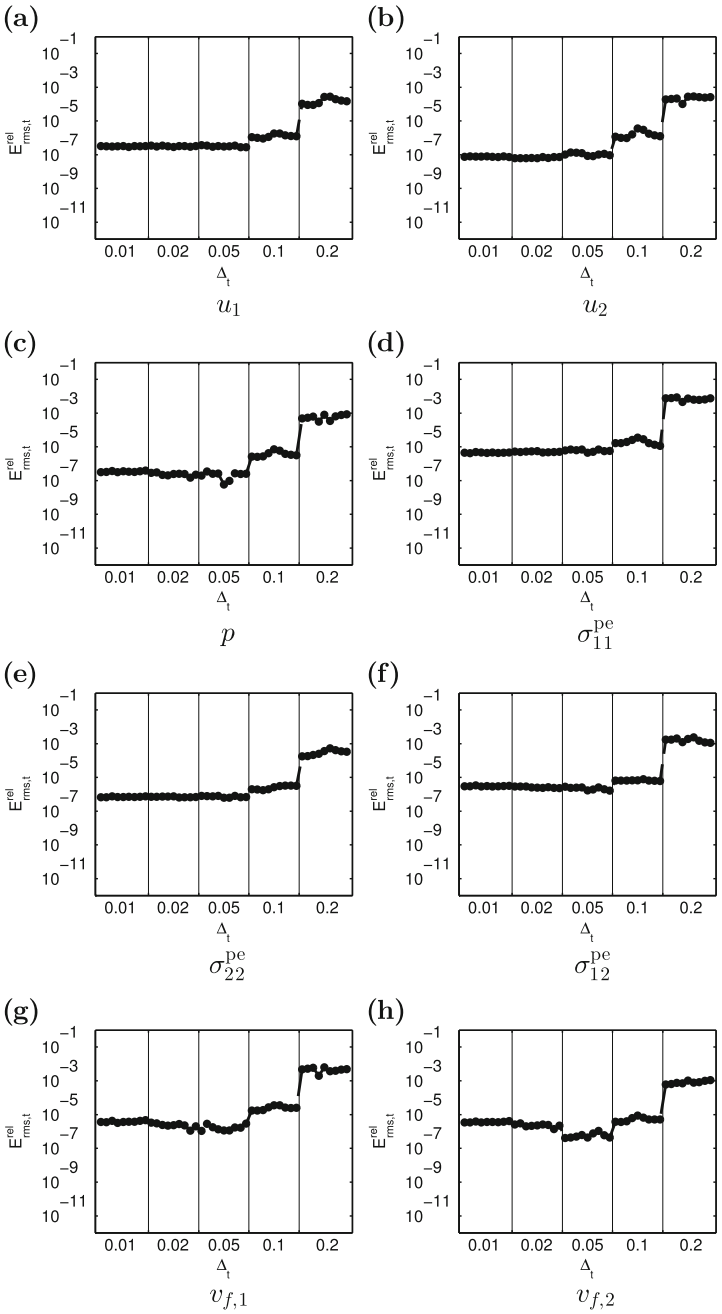

Fig. 6.31 Maximal relative rooted mean square errors $E^{rel}_{rms,t}$ as defined in (6.15) for Example 6.5 for varying parameters Δ_t, τ_1 as given in (6.18) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t$, $\tau_{-1} = 0.5\Delta_t$, $t_{end} = 1$. Each plot is split into five parts with the corresponding Δ_t given on the abscissa. For each part, τ_1 varies from small to large from left to right

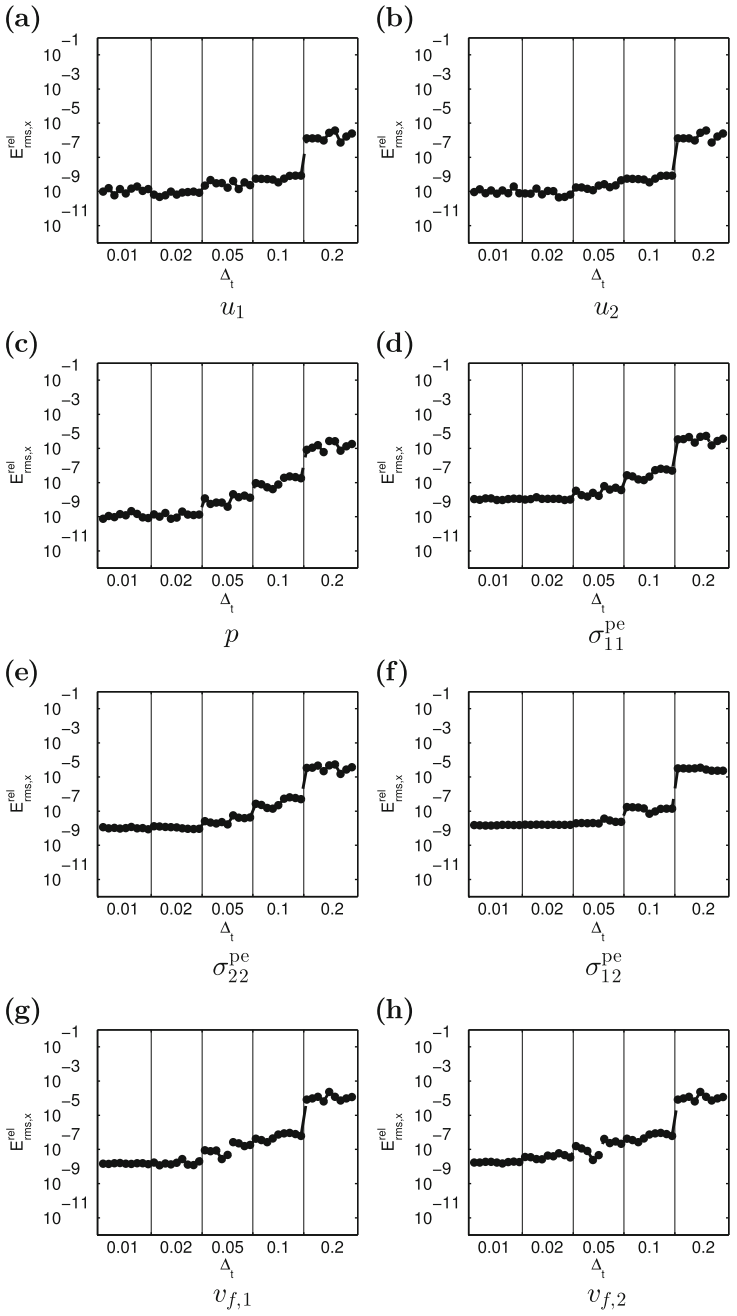

Fig. 6.32 Maximal relative rooted mean square errors $E^{rel}_{rms,x}$ as defined in (6.16) for Example 6.2 for varying parameters Δ_t, τ_1 as given in (6.18) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t$, $\tau_{-1} = 0.5\Delta_t$, $t_{end} = 1$. Each plot is split into five parts with the corresponding Δ_t given on the abscissa. For each part, τ_1 varies from small to large from left to right

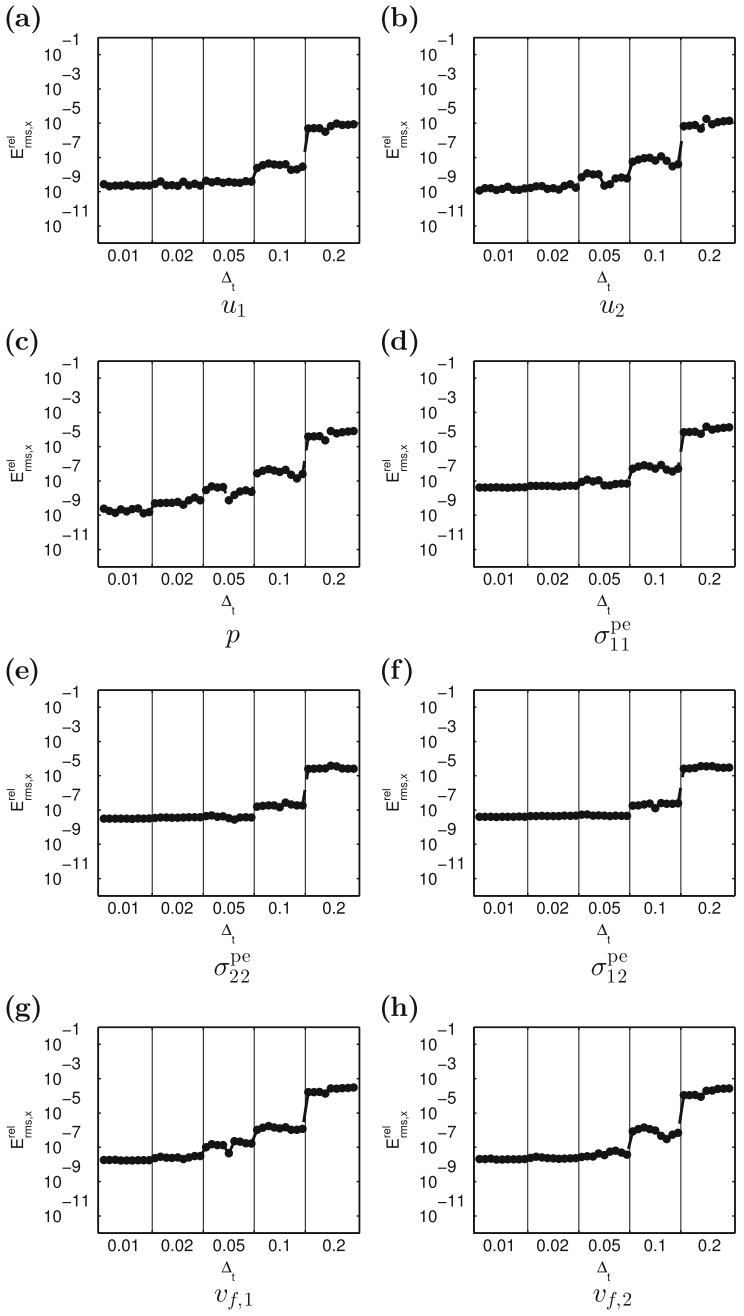

Fig. 6.33 Maximal relative rooted mean square errors $E^{rel}_{rms,x}$ as defined in (6.16) for Example 6.3 for varying parameters Δ_t, τ_1 as given in (6.18) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t$, $\tau_{-1} = 0.5\Delta_t$, $t_{end} = 1$. Each plot is split into five parts with the corresponding Δ_t given on the abscissa. For each part, τ_1 varies from small to large from left to right

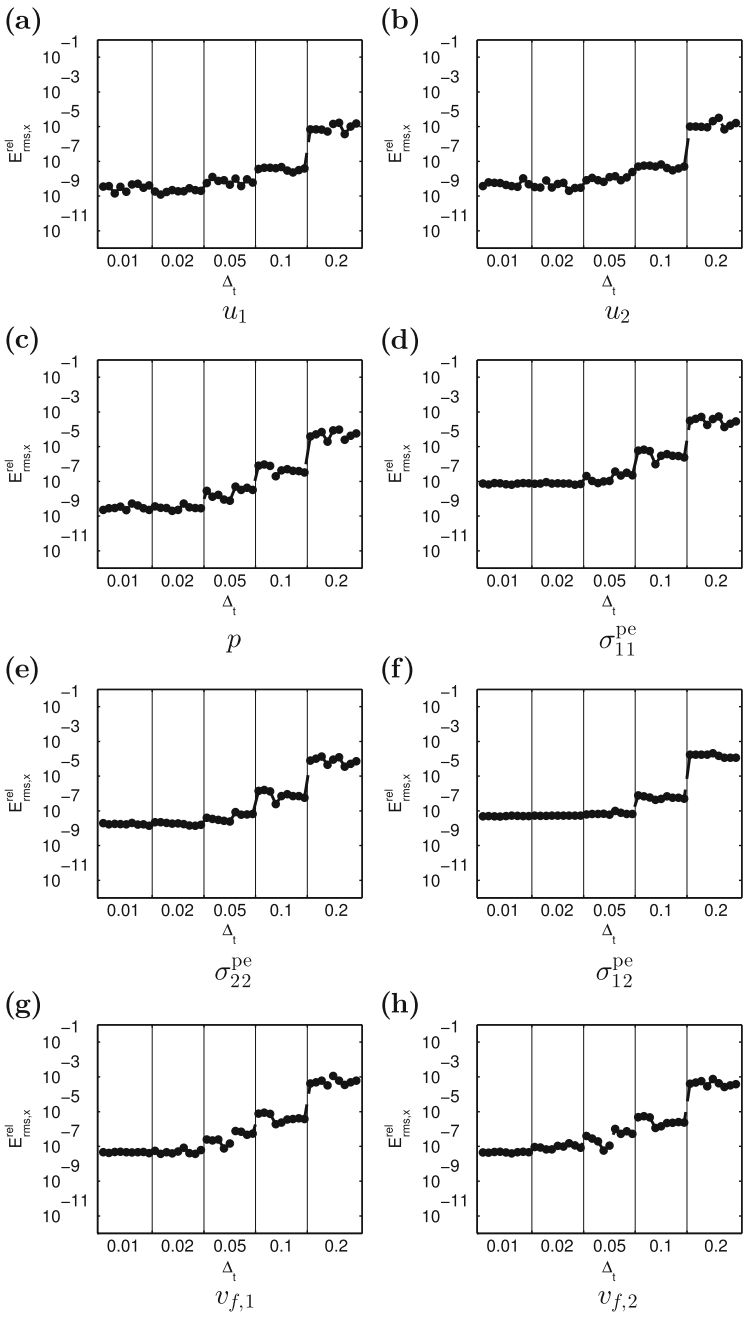

Fig. 6.34 Maximal relative rooted mean square errors $E^{rel}_{rms,x}$ as defined in (6.16) for Example 6.4 for varying parameters Δ_t, τ_1 as given in (6.18) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t$, $\tau_{-1} = 0.5\Delta_t$, $t_{end} = 1$. Each plot is split into five parts with the corresponding Δ_t given on the abscissa. For each part, τ_1 varies from small to large from left to right

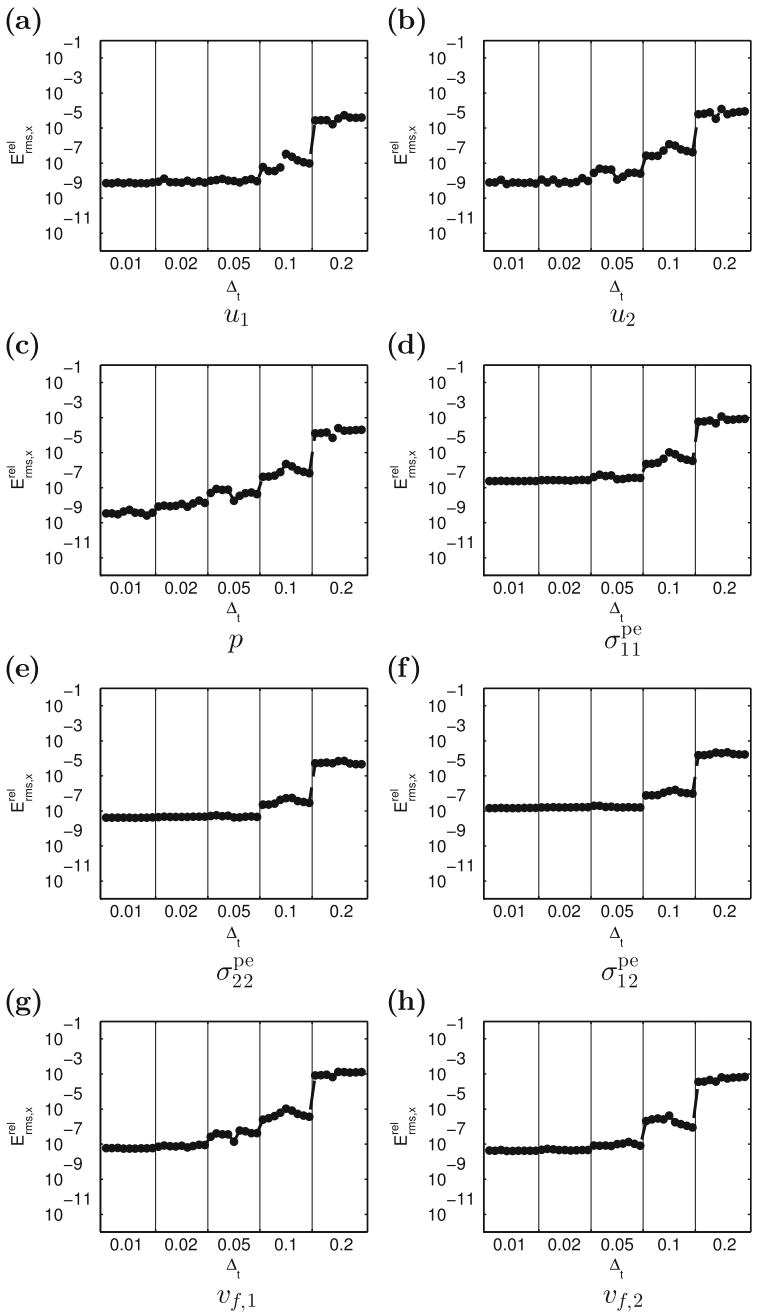

Fig. 6.35 Maximal relative rooted mean square errors $E^{rel}_{rms,x}$ as defined in (6.16) for Example 6.5 for varying parameters Δ_t, τ_1 as given in (6.18) and fixed values $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t$, $\tau_{-1} = 0.5\Delta_t$, $t_{end} = 1$. Each plot is split into five parts with the corresponding Δ_t given on the abscissa. For each part, τ_1 varies from small to large from left to right

Remark 6.6 It is, of course, possible to consider other configurations of source points. For example, all source points for Si-parts may be chosen to have negative τ. This would probably yield the same problems as choosing Δ_τ very small or unreasonably large, as the columns of the resulting matrix would be more similar than with the configuration chosen here. As there is no argument known to us why our method may benefit from another ratio of source points for Si-parts with positive τ to source points for Si-parts with negative τ, we do not investigate such choices here.

Remark 6.7 The reader familiar with numerical solution methods for partial differential equations which include time-dependence may wonder whether the increase in errors for large Δ_t may be due to violation of a CFL-condition [59, 60]. We assume that this is not the case. As we are not using any kind of time-marching or finite difference scheme with respect to time. As finite difference schemes are not the topic of this thesis, we did not investigate how a CFL-condition would limit the freedom of the choice for Δ_t. However, if we consider to first apply a finite difference scheme with respect to time and use the MFS for the resulting set of equations, we have to investigate the CFL-condition for quasistatic poroelasticity more closely.

6.2.4 Error Distribution in Time and Space

In the previous sections, we examined the dependence of numerical errors on the choice of parameters. It is also of interest to know how errors develop in time and how they are distributed in the domain. For this purpose, we now fix all parameters at the following values: $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$, corresponding to $I = 6$, $M = 4$, and $J = 100$.

Please note that the term "distribution" is used here in its everyday sense of how something is distributed, not as in the definition of a distribution as a generalized function.

Figure 6.36 gives the time development of the relative error $\mathrm{E}^{\text{rel}}(x = 0, t)$ for various poroelastic quantities at the origin, i.e., the center of Ω for Example 6.2. Figures 6.37 and 6.38 give the distribution of the relative error $\mathrm{E}^{\text{rel}}(x, t = 1)$ or $\mathrm{E}^{\text{rel}}(x, t = 5)$ on $\Omega = (-1, 1)^2$, respectively, for those quantities for the same example in a logarithmic pseudocolor plot using logarithmic contour distances. Figures 6.39–6.47 show the corresponding results for Examples 6.3–6.5, respectively. All figures depicting $\mathrm{E}^{\text{rel}}(x = 0, t)$ use the same scale, and those figures depicting $\mathrm{E}^{\text{rel}}(x, t = 1)$ or $\mathrm{E}^{\text{rel}}(x, t = 5)$ also use the same colorbar, for better comparability of the approximation error of the different poroelastic quantities within the same example as well as the same quantity between different examples. This is done to illuminate the influence of the function that is considered and the influence of the different boundary conditions we used, respectively.

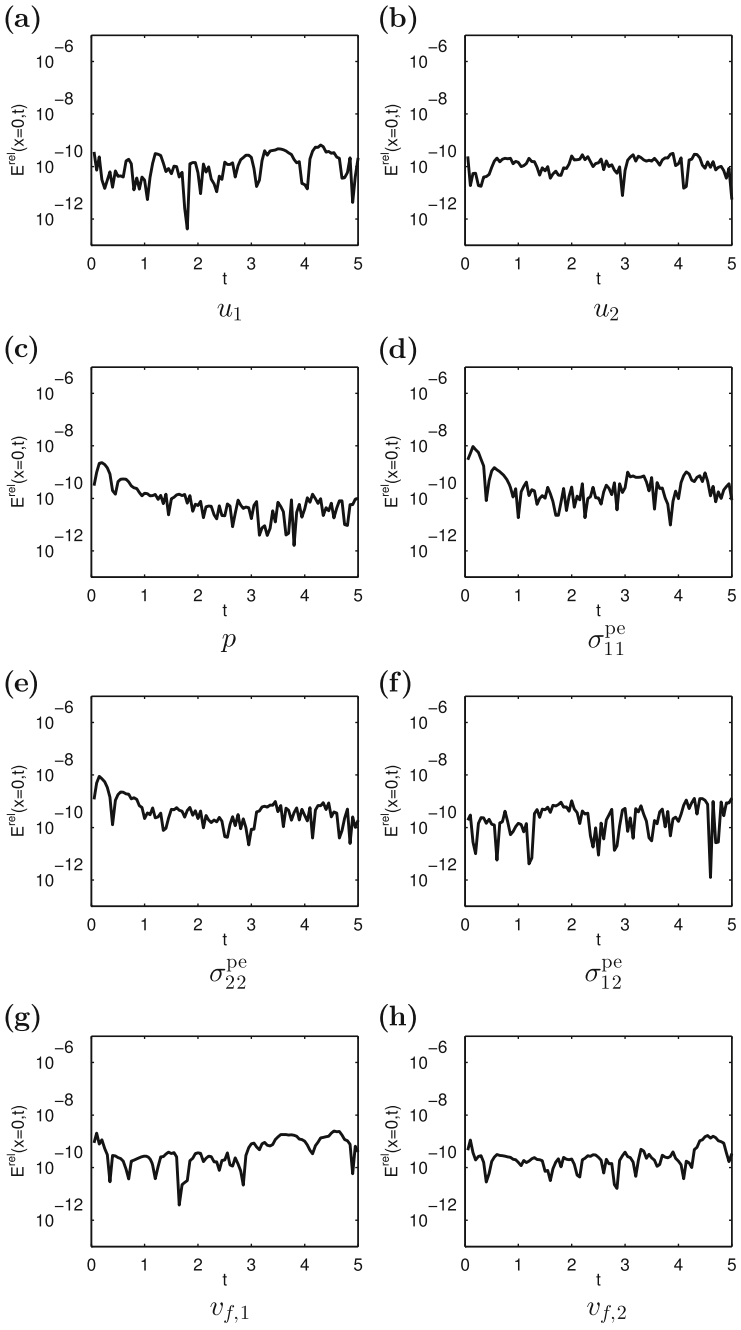

Fig. 6.36 Time development of the relative errors for approximating the poroelastic quantities for Example 6.2 evaluated at the origin $x = 0$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

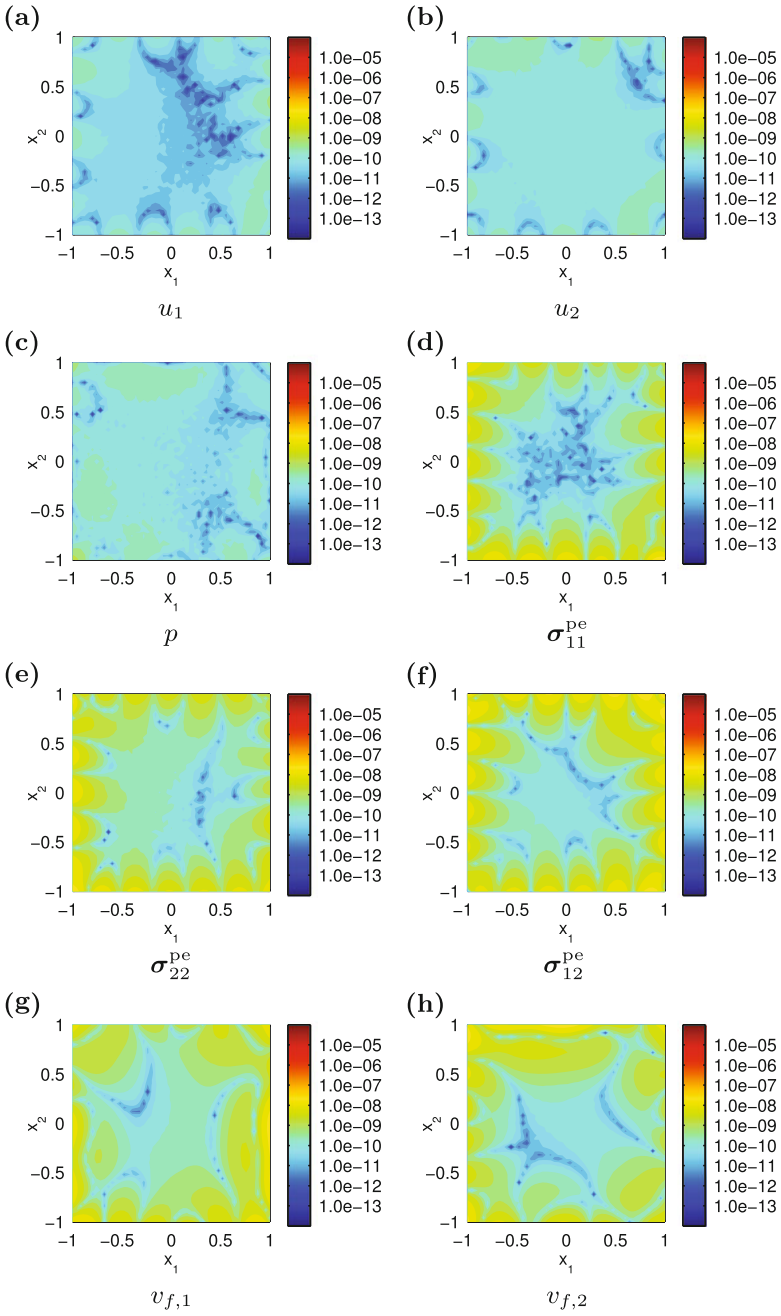

Fig. 6.37 Logarithmic pseudocolor plots using logarithmic contour distances for the relative errors for approximating the poroelastic quantities for Example 6.2 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$

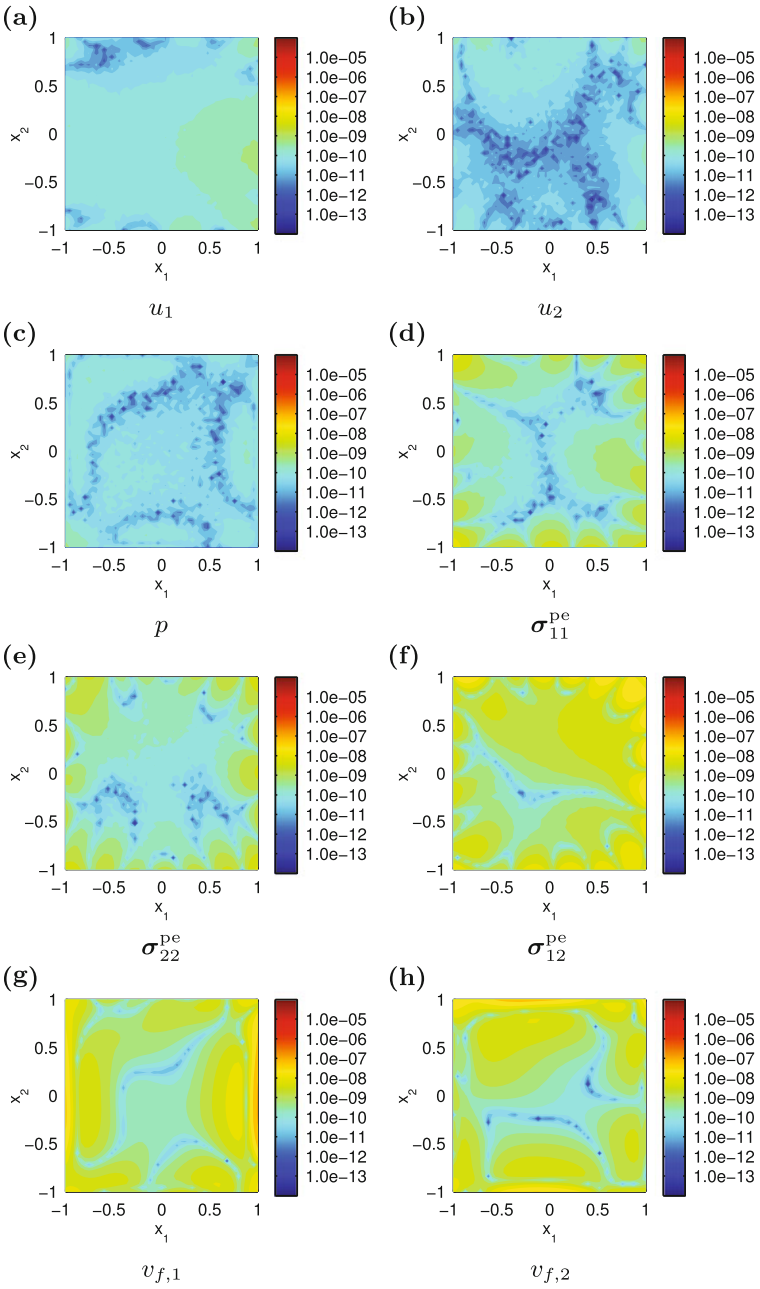

Fig. 6.38 Logarithmic pseudocolor plots using logarithmic contour distances for the relative errors for approximating the poroelastic quantities for Example 6.2 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$. In the little white spot in subfigure **b**, the error is too small to be captured by the colormap that is used

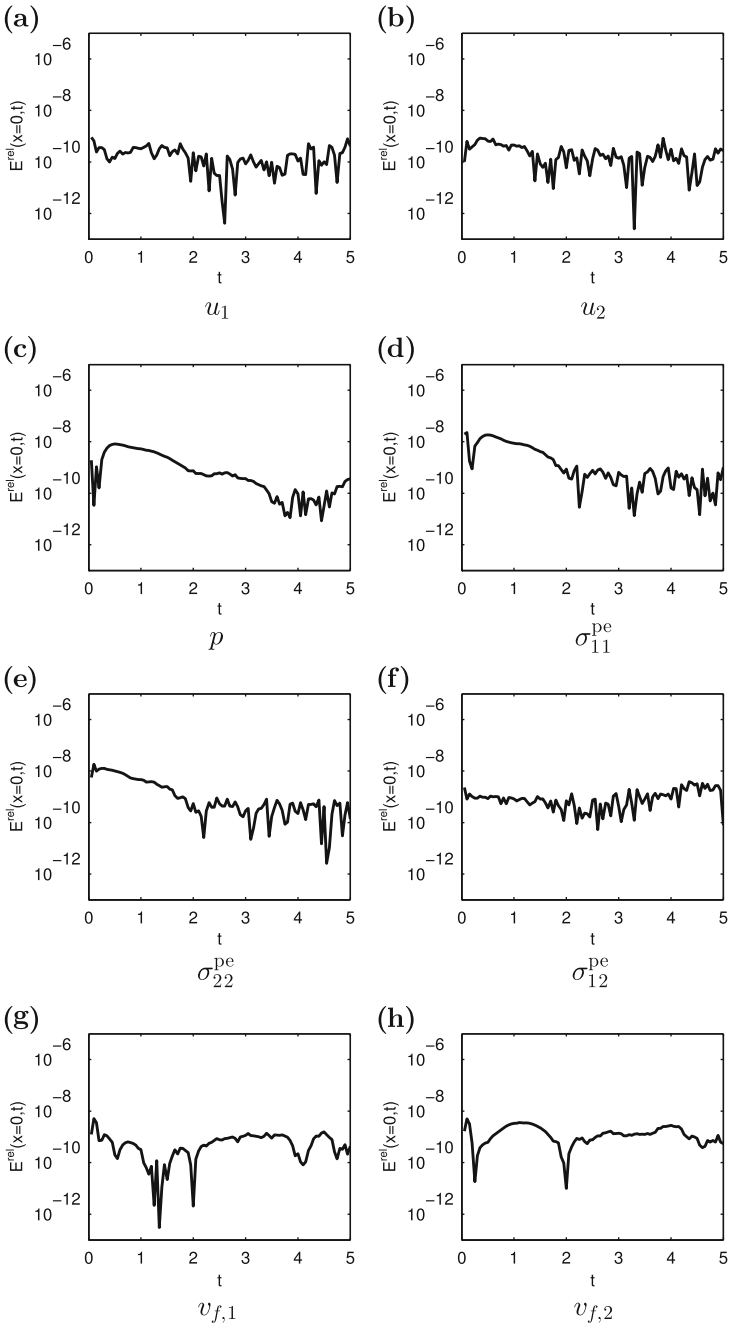

Fig. 6.39 Time development of the relative errors for approximating the poroelastic quantities for Example 6.3 evaluated at the origin $x = 0$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

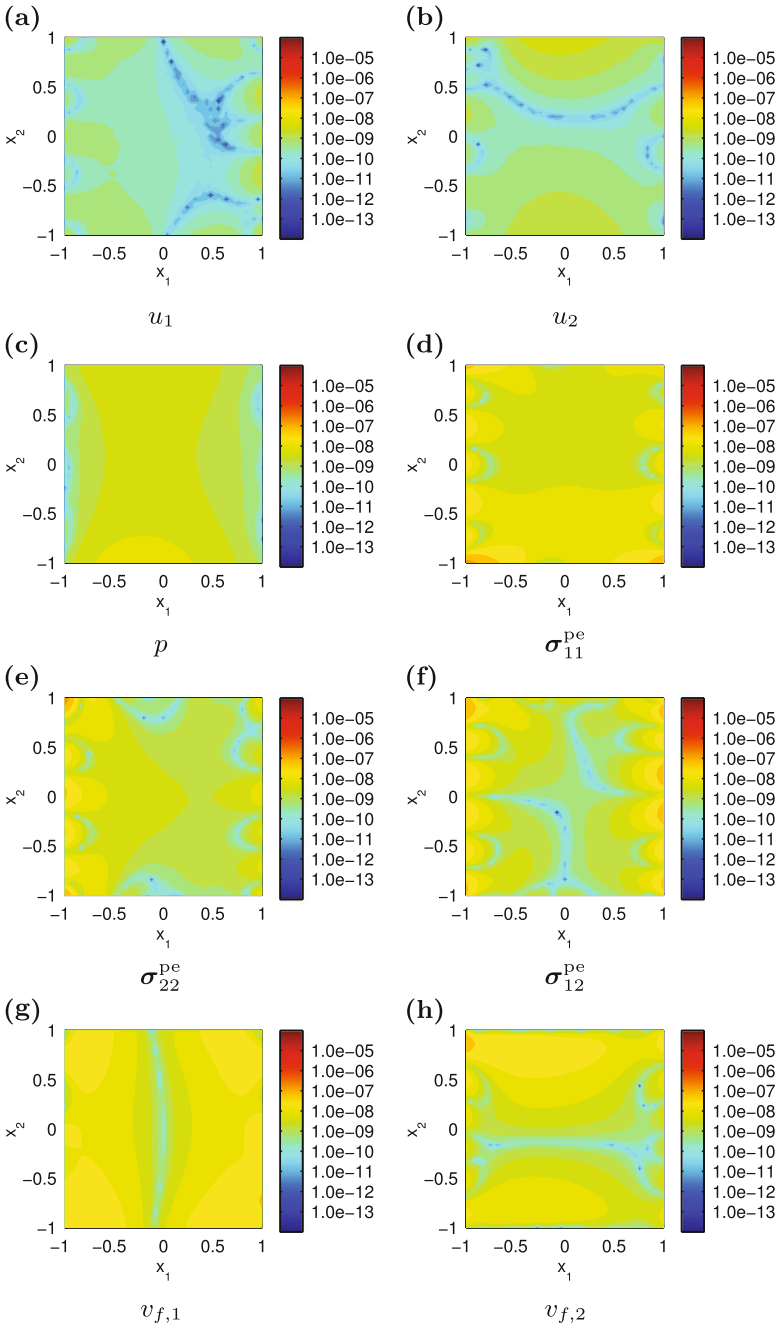

Fig. 6.40 Logarithmic pseudocolor plots of the relative errors for approximating the poroelastic quantities for Example 6.3 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

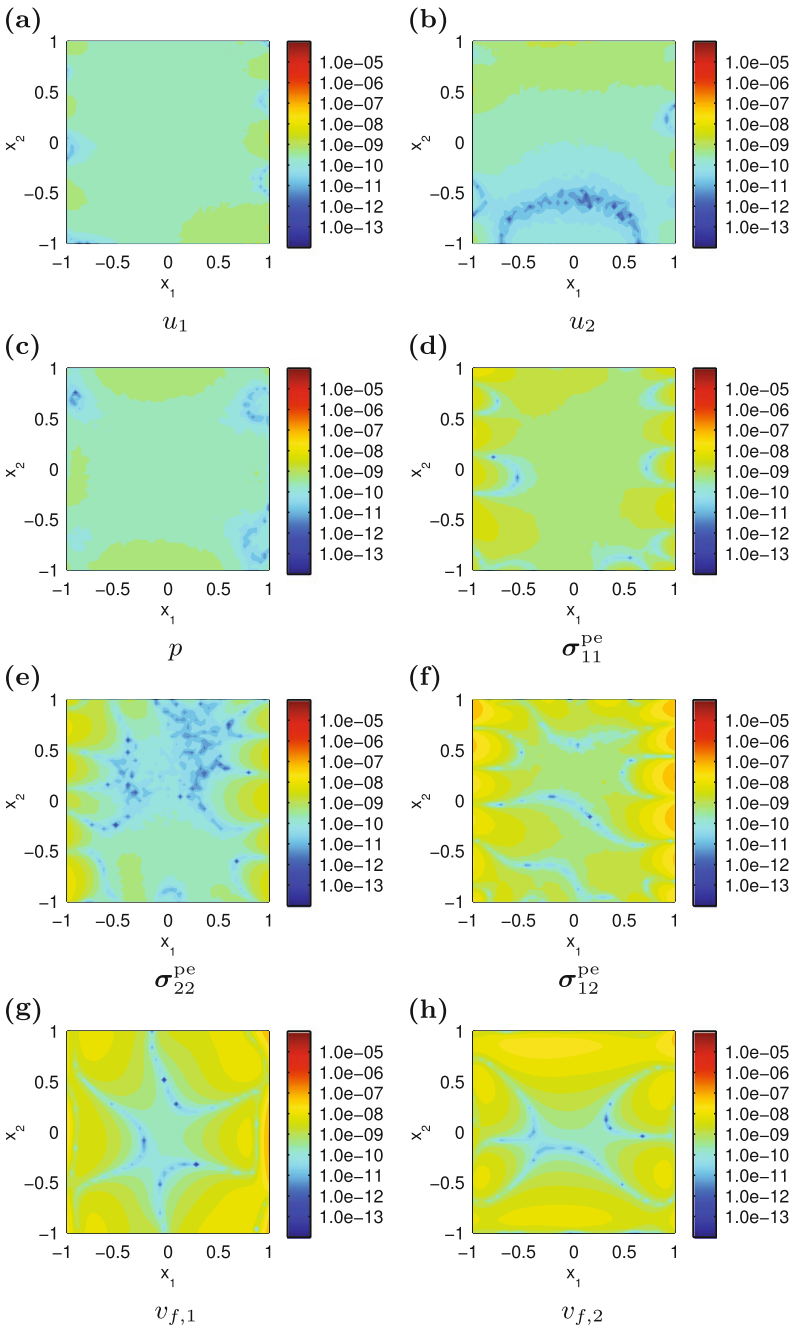

Fig. 6.41 Logarithmic pseudocolor plots of the relative errors for approximating the poroelastic quantities for Example 6.3 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

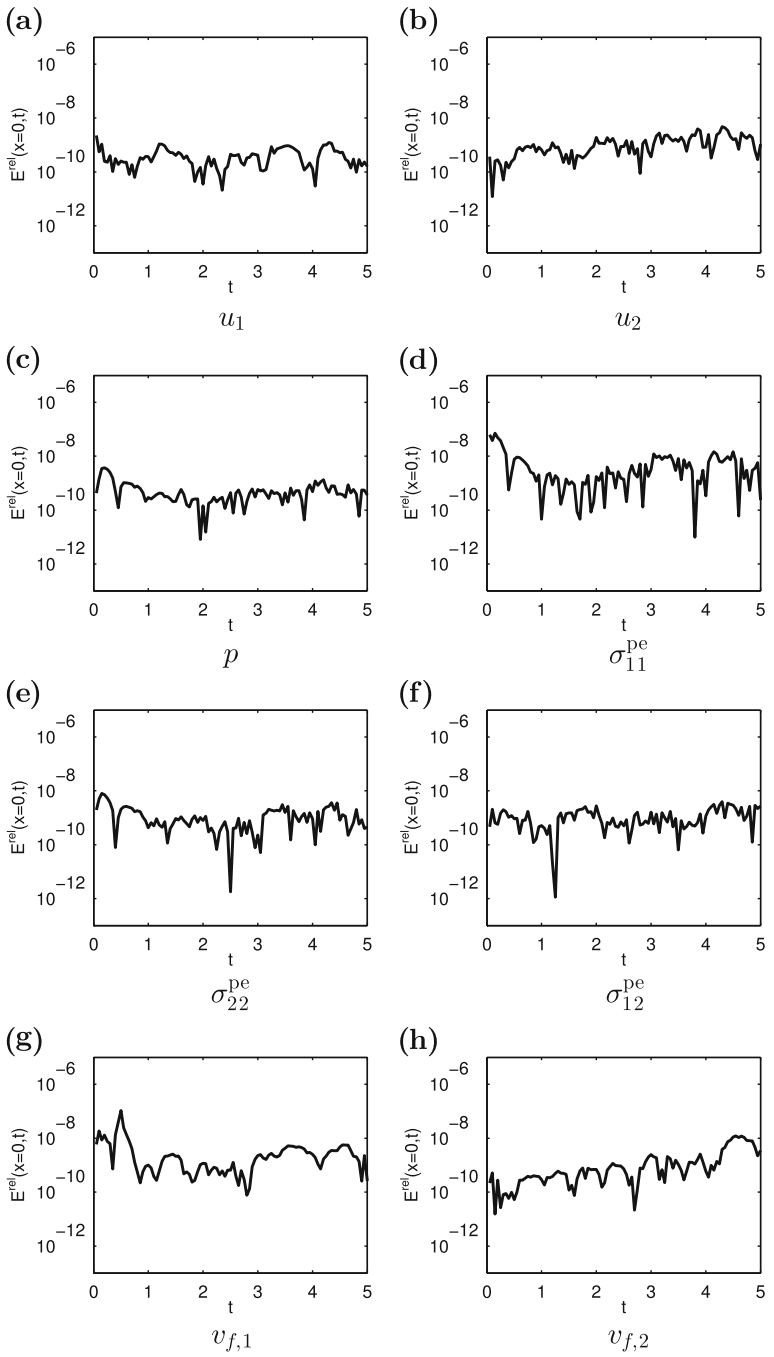

Fig. 6.42 Time development of the relative errors for approximating the poroelastic quantities for Example 6.4 evaluated at the origin $x = 0$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

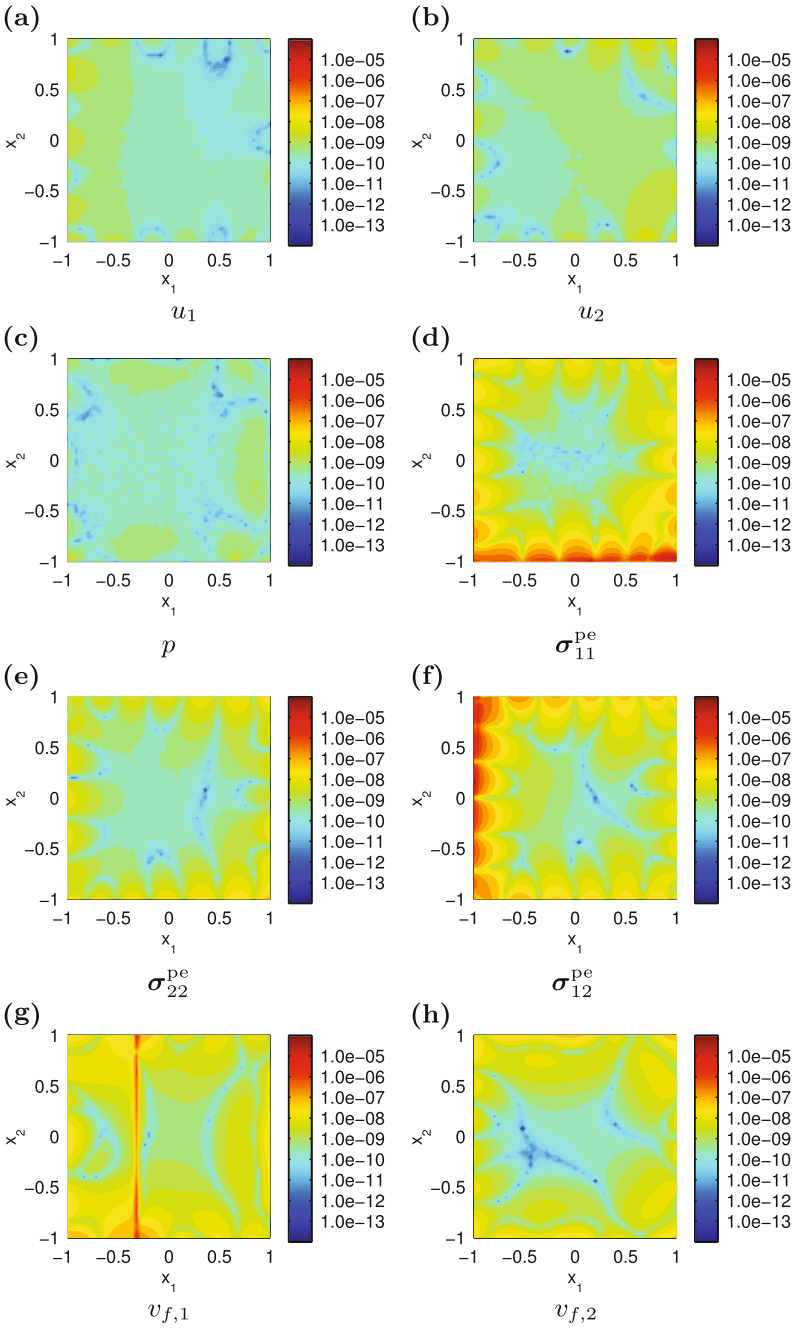

Fig. 6.43 Logarithmic pseudocolor plots of the relative errors for approximating the poroelastic quantities for Example 6.4 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

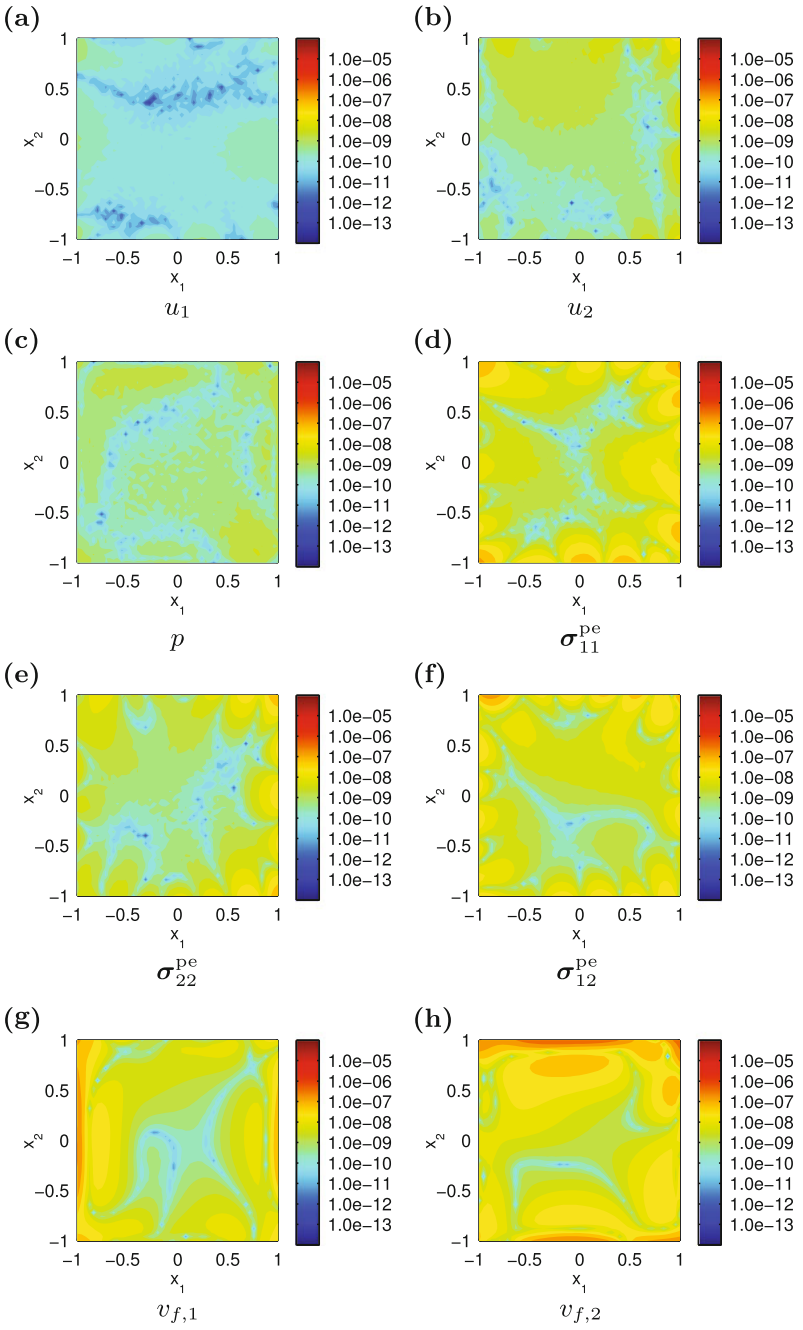

Fig. 6.44 Logarithmic pseudocolor plots of the relative errors for approximating the poroelastic quantities for Example 6.4 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

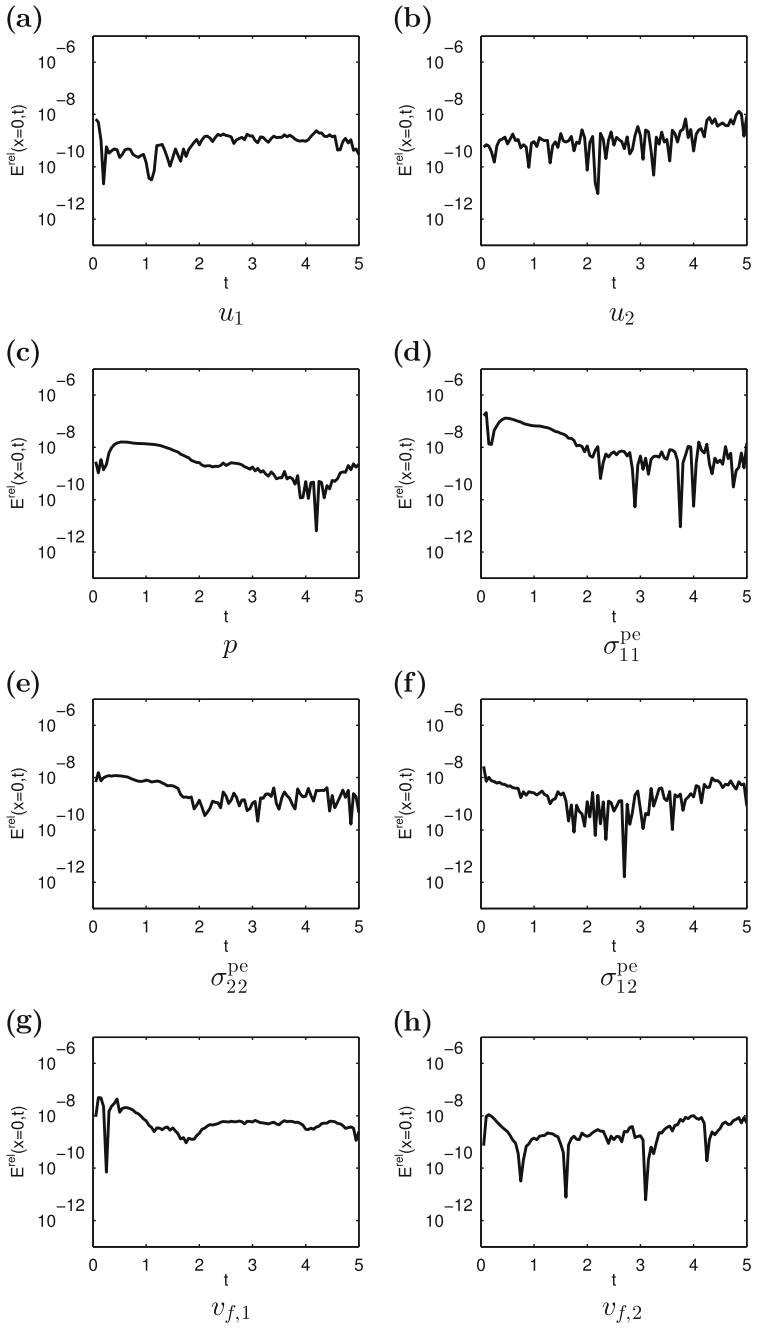

Fig. 6.45 Time development of the relative errors for approximating the poroelastic quantities for Example 6.5 evaluated at the origin $x = 0$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

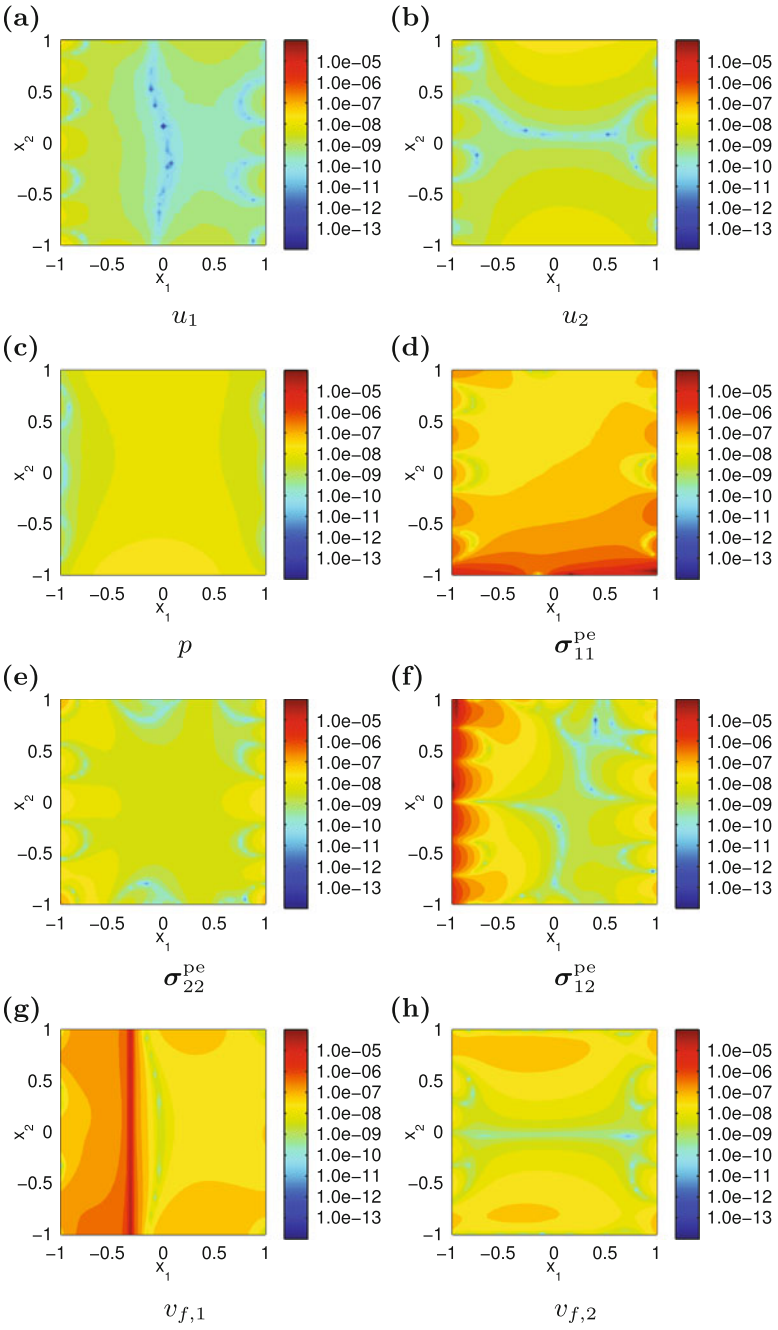

Fig. 6.46 Logarithmic pseudocolor plots of the relative errors for approximating the poroelastic quantities for Example 6.5 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

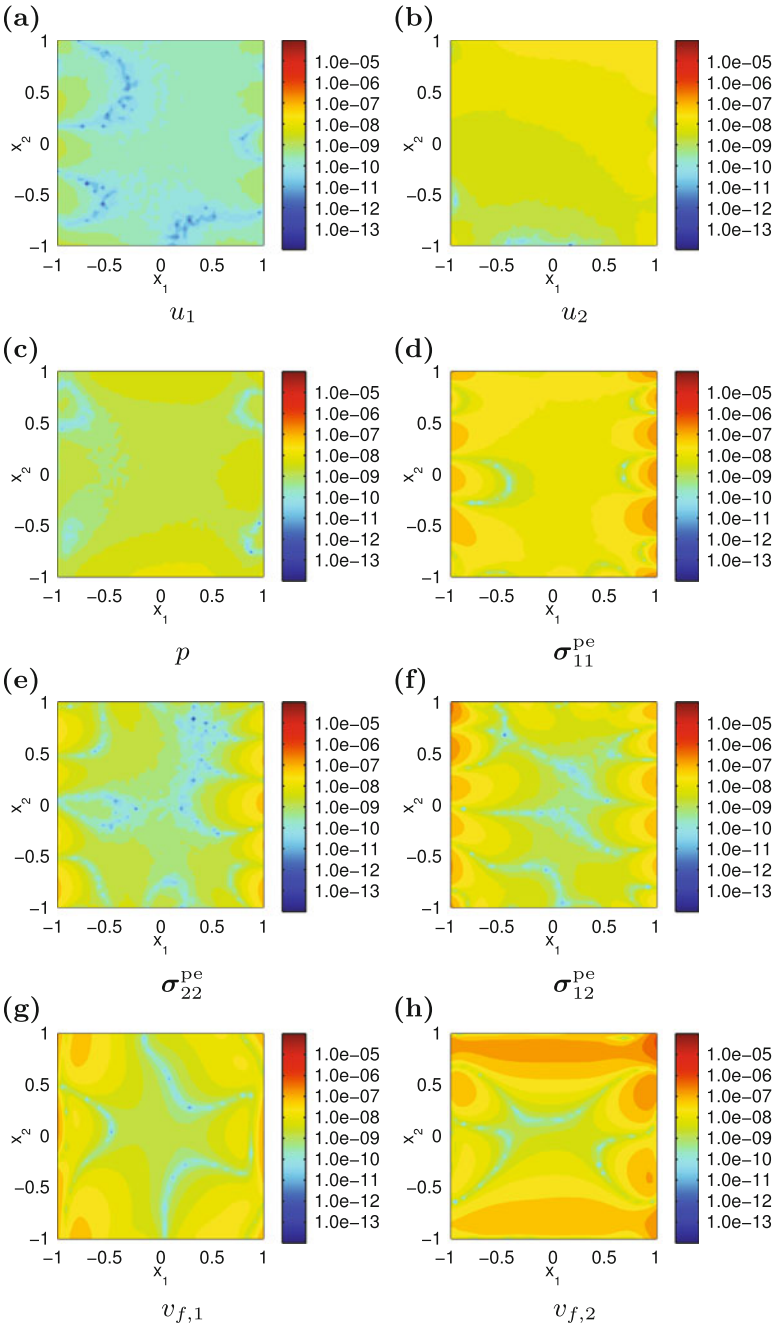

Fig. 6.47 Logarithmic pseudocolor plots of the relative errors for approximating the poroelastic quantities for Example 6.5 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\mathrm{end}} = 5$

First, we discuss the error distribution in time as given by Figs. 6.36, 6.39, 6.42, and 6.45. The behavior of the errors in the components of the displacement vector u cannot be described in a simple way as it changes rapidly and unsystematically between local maxima and minima. In contrast to this, the errors in the pressure p and the components of the stress tensor $\boldsymbol{\sigma}^{\mathrm{pe}}$ often show rather large errors for times near $t = 0$, which decrease with increasing time, especially when mixed boundary conditions are used (subfigures c–f in Figs. 6.39 and 6.45). For larger times, the error distributions lose this monotone behavior, looking more like the ones for u. The fluid velocity v_f shows a behavior similar to u if Dirichlet boundary conditions are considered (Figs. 6.36 and 6.42). In case of mixed boundary conditions, both components of v_f have some strong local minima in time and are almost constant for times between those minima (subfigures g–h in Figs. 6.39 and 6.45).

Considering the spatial distribution of errors, we notice a little white spot in Fig. 6.38b. In this spot, the error is too small to be captured by the colormap used here, i.e., it is smaller than 10^{-14}. We decided not to extend the colormap to provide a color for this spot as this would blur the boundaries between the other error levels too much.

In almost all figures and subfigures depicting relative errors in the domain for a fixed time, the largest errors are found at the boundary, usually in the corners. Although this is a common result, it may be a little bit surprising here as we primarily consider data at the boundary but not in the domain. It is in some sense in accordance with the fact that the largest errors are also found for small time, i.e., close to the prescribed initial condition. A possible interpretation of these results is that inside of the domain and as time progresses, solutions to the QEP tend towards a similar behavior such that the influence of initial conditions is reduced over time and the influence of boundary conditions reduces with increasing distance from the boundary. This is supported by results on exponential damping in time known for solutions to parabolic equations [101]. However, as Eqs. (4.1) are not purely parabolic, further, more specific investigations are needed to prove results on the general behavior of solutions.

An exception from the statement that the largest errors are found at the boundary is $v_{f,1}$ for the examples considering an fi-solution as can be seen in Figs. 6.43h and 6.46h. Here, the largest errors occur in a strip almost parallel to the x_2-axis. This can be explained by a comparison to Fig. 6.8h, which shows that $v_{f,1}$ almost vanishes within this strip, such that the large relative errors are due to the small absolute values of $v_{f,1}$ in the denominator of $\mathrm{E}^{\mathrm{rel}}(x, t)$ (see (6.10a)).

All in all, errors in $\boldsymbol{\sigma}^{\mathrm{pe}}$ and v_f are larger than errors in u and p, which are our primary variables. Moreover, in agreement with the results of the previous chapters, the errors obtained with mixed boundary conditions are larger than the ones we get if pure Dirichlet boundary conditions are prescribed. Furthermore, errors for Examples 6.4 and 6.5 are larger than those for Examples 6.2 and 6.3, which also agrees with our previous results.

Comparing the error distributions for $t = 1$ and $t = 5$ yields different results for each example. For Example 6.2, the maximum errors in u_1, $\boldsymbol{\sigma}^{\mathrm{pe}}_{11}$, $v_{f,1}$, and $v_{f,2}$ increase with increasing time, whereas they decrease for all other

quantities. However, as all errors in this example are very small, those changes are probably negligible. The same statements are true with regard to Example 6.3. For Examples 6.4 and 6.5, we see that the errors in u_2 and $v_{f,2}$ are increasing, whereas the errors in the other quantities are decreasing in time. However, the reader should not be misled by these results as we only consider two snapshots with respect to time. As discussed above, the distribution of errors in time is not monotone. A correct analysis of the development of the errors in the whole domain over time can only be based on theoretical results, i.e., theorems about convergence of the MFS. As mentioned before, such an analysis is not yet available, even for less complicated equations, as, e.g., the heat equation.

To put our results into context, we compare them with some results available in the literature. Probably the closest to our method is the work by Wen and Liu [288], who considered an MFS for dynamic poroelasticity in the Laplace domain. In their paper, they report relative mean errors of 10^{-2}–10^{-3}. Finite element methods for quasistatic poroelasticity are regarded in [221–225, 293, 294]. The best errors reported in these articles and PhD-theses are between 10^{-3} and 10^{-6}. As we can see, our method of fundamental solutions can compete with those finite element methods.

6.3 Using a Gaussian Least-Squares Method vs. Using SVD

Although the singular value decomposition comes with the advantage of being robust and is, thus, well suited for ill-conditioned systems, its requirements in terms of memory and computational time are rather high, as both matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ as well as the singular values have to be computed. Computations could be accelerated and reduced in memory requirements if the system of linear equations could be solved by simple Gaussian elimination, or more correctly, a simple least-squares method. Both are implemented in Matlab by the backslash operator [190], with the latter being used automatically if the system under consideration is not square. As we did not use any regularization technique for the singular value decomposition, we do also not apply any regularization to the Gaussian least-squares method.

In this section, we use the backslash operator instead of the svd routine to compute solutions to Examples 6.3 and 6.5, i.e., those with mixed boundary conditions. As we already discussed the approximation error in detail, we only consider one parameter combination here: $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$, corresponding to $I = 6$, $M = 4$, and $J = 100$. This is the same as in Sect. 6.2.4. Consequently, the errors are not given graphically but in Tables 6.4 and 6.5. As can be seen from these tables, all errors are still very small. In fact, they are at most increased by one order of magnitude compared to the ones achieved by using the singular value decomposition. In particular, the errors that are an order of magnitude larger are $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}$ for $\boldsymbol{\sigma}^{\mathrm{pe}}_{11}$ and $\boldsymbol{\sigma}^{\mathrm{pe}}_{22}$ in Example 6.3 and $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}$ for p and $v_{f,2}$ in the same example. Considering that all errors are still very small, this differences are negligible, such that a simple

Table 6.4 Errors in the poroelastic quantities for Example 6.3 using a Gaussian least-squares method. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as defined in (6.15) and (6.16) and the point $x \in \mathcal{I}_2$ or time $t \in \mathcal{I}_t$ at which they are obtained

Quantity	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}$	x_1	x_2	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t
u_1	7.8838e–08	−1.0	1.0	1.8598e–09	1.0
u_2	9.5941e–08	1.0	1.0	8.1788e–09	1.0
p	6.2620e–08	1.0	−1.0	1.3340e–08	1.0
$\boldsymbol{\sigma}^{\text{pe}}_{11}$	5.2914e–06	1.0	−1.0	9.6958e–08	1.0
$\boldsymbol{\sigma}^{\text{pe}}_{22}$	5.2433e–07	1.0	−1.0	1.5273e–08	1.0
$\boldsymbol{\sigma}^{\text{pe}}_{12}$	1.6671e–06	−1.0	1.0	4.6767e–08	1.0
$v_{f,1}$	7.3198e–07	1.0	0.0	7.5549e–08	1.0
$v_{f,2}$	6.1712e–07	1.0	1.0	1.4920e–07	5.0

Table 6.5 Errors in the poroelastic quantities for Example 6.5 using a Gaussian least-squares method. Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as defined in (6.15) and (6.16) and the point $x \in \mathcal{I}_2$ or time $t \in \mathcal{I}_t$ at which they are obtained

Quantity	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}$	x_1	x_2	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t
u_1	2.1582e–08	1.0	1.0	5.9641e–10	1.0
u_2	1.8411e–08	−1.0	1.0	1.3295e–09	1.0
p	1.3365e–08	0.0	−1.0	4.8341e–09	1.0
$\boldsymbol{\sigma}^{\text{pe}}_{11}$	3.4114e–07	1.0	−1.0	1.3351e–08	1.0
$\boldsymbol{\sigma}^{\text{pe}}_{22}$	4.4697e–07	−1.0	1.0	8.7836e–09	1.0
$\boldsymbol{\sigma}^{\text{pe}}_{12}$	2.3679e–07	1.0	0.5	1.6251e–08	5.0
$v_{f,1}$	5.3512e–07	1.0	0.0	2.6854e–08	5.0
$v_{f,2}$	1.9527e–07	−1.0	1.0	2.1262e–08	5.0

Gaussian least-squares algorithm can be used as a solution method to the system of linear equations which we encounter. This may be surprising as the spectral condition number of the system matrix is $7.5 \cdot 10^{16}$. However, we consider here only examples for which the right-hand side is known up to machine precision. Considering the influence of noise is out of the scope of this thesis, but an important topic of further research. The interested reader is referred to [212] for a surface impression on how regularization may be achieved and to [29, 30] and the references therein for exemplary applications of two regularization methods considering either the singular value decomposition or Gaussian elimination for ill-posed problems with noise.

6.4 Comparison with the Ansatz Including fi-Parts

As our first derivation of an MFS led to Ansatz (5.10) which includes fi-parts of the fundamental solutions, we implemented this ansatz, too, and make a brief comparison. The source points for fi-parts are the same as the source points for the Si-parts (cf. Fig. 6.1). Changes in the system matrix appear in the submatrices $\left(\boldsymbol{A}_{kl}^{(1)}\right)_{k=3(i-1)+1,l=3(m-1)+1}^{3i,3m}$ and $\left(\boldsymbol{A}_{kl}^{(j)}\right)_{k=3(i-1)+1,l=3(m-1)+1}^{3i,3m}$ given by (6.7d) and (6.7e). To each submatrix of those types, two additional columns are added for each source point, containing the contributions of fi-parts. Those contributions are given by

$$\begin{pmatrix} \boldsymbol{u}_{11}^{\mathrm{fi}}(x_{(i)}-y_{(m)},t_j-\tau_1) & \boldsymbol{u}_{12}^{\mathrm{fi}}(x_{(i)}-y_{(m)},t_j-\tau_1) \\ \boldsymbol{u}_{21}^{\mathrm{fi}}(x_{(i)}-y_{(m)},t_j-\tau_1) & \boldsymbol{u}_{22}^{\mathrm{fi}}(x_{(i)}-y_{(m)},t_j-\tau_1) \\ p_1^{\mathrm{fi}}(x_{(i)}-y_{(m)},t_j-\tau_1) & p_2^{\mathrm{fi}}(x_{(i)}-y_{(m)},t_j-\tau_1) \end{pmatrix}$$

$$i,j,m \in \mathbb{N}, i \le 4I, 1 \le j \le J, m \le 4M\ . \tag{6.20}$$

For normal tensions and fluid velocities, the corresponding contributions are given by the poroelastic stress tensors and fluid velocities related to the fi-parts as given in Appendix A. The resulting system of linear equations is solved using the svd routine to allow for a fair comparison to our previous results.

As before, we restrict our analysis to Examples 6.3 and 6.5 and fix $\Delta_x = \frac{1}{3}$, $\Delta_t = \Delta_\tau = 0.05$, $\tau_1 = \tau_{-1} = 0.5\Delta_t$, and $t_{\mathrm{end}} = 5$. As we have seen in Sect. 6.2.1, Δ_y and γ have a significant influence on the performance of the MFS. Thus, different combinations according to (6.14) are tested. Here, we report only the best results to see if accuracy is lost by using the reduced Ansatz (6.5) in Tables 6.6 and 6.7.

As we can see, the best results are obtained for $\Delta_y = 0.2$ ($M = 10$), i.e., an underdetermined system. This corresponds to the results in [267] for the Laplace equation. The condition number in this case is $1.7 \cdot 10^{17}$, roughly twice as large as before. This may explain why although the computational effort is much higher, the errors are roughly an order of magnitude larger on average. We should not be surprised by this result as we have added more rather similar columns to the system matrix. Even for Example 6.5, the results are not improved although the full Ansatz (5.10) includes the fi-parts of the fundamental solutions, which may be expected to improve the approximation of another fi-solution.

All in all, we can conclude that the usage of the reduced Ansatz (6.5) is preferable.

Table 6.6 Errors in the poroelastic quantities for Example 6.3 using Ansatz (5.10). Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{5}$ corresponding to $M = 10$, γ as given in the table, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as defined in (6.15) and (6.16) and the point $x \in \mathcal{I}_2$ or time $t \in \mathcal{I}_t$ at which they are obtained

Quantity	M	γ	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}$	x_1	x_2	γ	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t
u_1	10	3.0	1.8172e–07	1.0	1.0	3.0	1.9319e–07	5.0
u_2	10	3.0	1.3797e–07	−1.0	1.0	3.0	6.2713e–08	5.0
p	10	3.0	4.6104e–07	−1.0	−1.0	3.0	5.4095e–07	5.0
$\boldsymbol{\sigma}^{\text{pe}}_{11}$	10	3.0	2.4919e–06	−1.0	−1.0	3.0	4.9722e–07	5.0
$\boldsymbol{\sigma}^{\text{pe}}_{22}$	10	3.0	3.5326e–06	−1.0	1.0	3.0	8.7850e–07	5.0
$\boldsymbol{\sigma}^{\text{pe}}_{12}$	10	3.0	6.3035e–06	1.0	1.0	3.0	2.4908e–06	5.0
$v_{f,1}$	10	3.0	9.0202e–06	1.0	0.0	3.0	3.9189e–05	5.0
$v_{f,2}$	10	3.0	1.0159e–05	−1.0	−1.0	3.0	8.5745e–06	5.0

Table 6.7 Errors in the poroelastic quantities for Example 6.5 using Ansatz (5.10). Parameters are $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{5}$ corresponding to $M = 10$, γ as given in the table, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as defined in (6.15) and (6.16) and the point $x \in \mathcal{I}_2$ or time $t \in \mathcal{I}_t$ at which they are obtained

Quantity	M	γ	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}$	x_1	x_2	γ	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t
u_1	10	3.0	6.5480e–07	−1.0	1.0	3.0	1.3498e–07	5.0
u_2	10	3.0	1.4941e–06	1.0	1.0	3.0	7.9817e–07	5.0
p	10	3.0	9.2556e–07	1.0	−1.0	3.0	2.6050e–06	5.0
$\boldsymbol{\sigma}^{\text{pe}}_{11}$	10	2.5	3.8625e–05	1.0	−1.0	3.0	6.1792e–06	5.0
$\boldsymbol{\sigma}^{\text{pe}}_{22}$	10	3.0	1.2993e–05	1.0	−1.0	2.5	2.1681e–06	5.0
$\boldsymbol{\sigma}^{\text{pe}}_{12}$	10	3.0	2.1083e–05	−1.0	1.0	2.5	3.5693e–06	5.0
$v_{f,1}$	10	3.0	2.5205e–05	−1.0	−1.0	2.5	4.0648e–05	5.0
$v_{f,2}$	10	3.0	1.0846e–05	1.0	1.0	3.0	4.7243e–05	5.0

6.5 A Time-Marching Scheme

In our previous choices of collocation and source points, we let ourselves be guided by Johansson et al. [143]. However, in [298, 300], Young et al. proposed a different approach, which has the characteristics of a time-marching scheme. The idea of this approach is as follows: we start with a given initial condition at $t = 0$ and boundary conditions for the first time step $t = \Delta_t$. With these data, a solution inside Ω is computed for $t = \Delta_t$. This solution is now considered as a new initial condition from which, together with boundary data for $t = \Delta_t$ and $t = 2\Delta_t$, a solution inside Ω for $t = 2\Delta_t$ is computed. This procedure is repeated until the final time t_{end} is reached.

Figure 6.48 shows the collocation and source points for the time-marching scheme according to [298, 300]. For a better overview, we include a separate plot for the collocation points (blue) at $t = 0$. The layout of the collocation points for $t = \Delta_t$

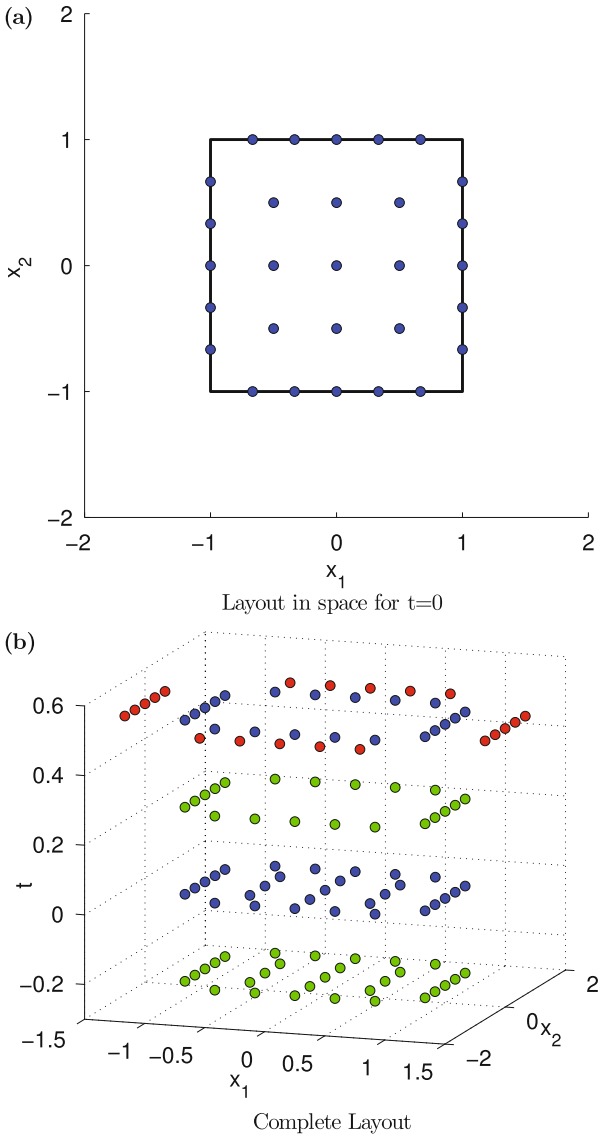

Fig. 6.48 Positions of the collocation points (*blue*) and source points (*red* for the CN- and St-parts, *green* for the Si-parts) used to compute a solution to the QEP (4.1) with a reduced MFS when using the time-marching scheme according to [298, 300]

only contains the points at the boundary in the same spatial location as for $t = 0$. The source points for the CN- and St-parts (red) are obtained by shifting the collocation points at the boundary in direction of the outer normal to Ω. They are only relevant for $t = \Delta_t$ as they don't contribute to ζ. The source points for the Si-parts (green) are chosen differently. For those source points at $t = -0.5\Delta_t$, their spatial coordinates are the same as for the collocation points at $t = 0$ (blue), including points inside Ω. For those source points at $t = 0.5\Delta_t$, their spatial coordinates are the same as for the collocation points at $t = \Delta_t$ (blue). This means that γ only influences the location of the source points for the CN- and St-parts (red). The source points for the Si-parts are not chosen on a pseudo-boundary, but on the boundary Γ and even in the domain Ω, but for different times. We have $\Delta_y = \Delta_x$. In a difference to Fig. 6.1, the points on the boundary which are nearest to the corners are given by setting one x-coordinate to ± 1 and the other to $\pm(1 - \Delta_x)$, following [298, 300], whereas in Fig. 6.1 it is set to $\pm\left(1 - \frac{\Delta_x}{2}\right)$.

A reasonable restriction on Δ_x is $\Delta_x^{-1} \in \mathbb{N}$, allowing the choice $\Delta_{x,0}^{-1} = \Delta_x^{-1} - 1$. Source points inside Ω are given by the cartesian grid

$$\left\{x_1 = -1 + i\Delta_{x,0}, i \in \mathbb{N}, i < \frac{2}{\Delta_{x,0}}\right\} \times \left\{x_2 = -1 + j\Delta_{x,0}, j \in \mathbb{N}, j < \frac{2}{\Delta_{x,0}}\right\} . \tag{6.21}$$

Figure 6.48 follows [300]. However, in [298], it is stated that better results are obtained if source points are not chosen at $t = \pm\Delta_t$, but at $t = (1 - \delta_{\mathrm{TM}})\Delta_t$ and $t = -\delta_{\mathrm{TM}}\Delta_t$ with $\delta_{\mathrm{TM}} \in \mathbb{R}^+$ a positive constant. For the above figure, we have $\delta_{\mathrm{TM}} = 0.5$. In [298], it is suggested to choose δ_{TM} equal to the diameter of Ω. In our case, this results in $\delta_{\mathrm{TM}} = \sqrt{8}$ for $\Omega = (-1, 1)^2$, whereas in [298], the domain $(0, 1)^2$ is considered, yielding $\delta_{\mathrm{TM}} = \sqrt{2}$. With this choice, Fig. 6.48 is misleading as all green source points are now chosen at a time prior to $t = 0$. Thus, both layers of the Si-parts contribute to the approximation of the initial condition.

The advantage of this time-marching scheme is that instead of one large system of linear equations, many small systems – one for each $t = j\Delta_t, j \in \mathbb{N}, 0 < j \leq J$ – with different right-hand sides have to be solved. As those systems are in most cases much smaller, even if Δ_x takes a rather small value, the computation time is significantly reduced.

As Prof. Dr. D.-L. Young and Prof. Dr. C.-M. Fan stated in personal communication, they implemented their MFS in FORTRAN, solving the resulting system of linear equations either with the LSARG routine from the IMSL library [283] or the LU decomposition from Numerical Recipes [228], and in Matlab [190] using the backslash operator. For the sake of comparability, we use the svd routine. However, as each system that we have to solve incorporates the same matrix, but different right-hand sides, the singular value decomposition only has to be computed once.

In order to become acquainted with the time-marching scheme, we tried to reproduce some results from [300]. Using the same parameters, the errors we

obtained for [300, Example 4.1] were about an order of magnitude larger. In this example, the function (adapted to our notation)

$$\zeta(x,t) = \left(\cos\left(\frac{\pi}{2}x_1\right) + \sin\left(\frac{\pi}{2}x_1\right) + \cos\left(\frac{\pi}{2}x_2\right) + \sin\left(\frac{\pi}{2}x_2\right)\right)\exp\left(-\frac{\pi^2}{4}t\right) \tag{6.22}$$

is considered on $\Omega = [0,1]^2$ for $t \in (0,3)$. The best we could get (using the backslash operator) was a relative error about $6 \cdot 10^{-4}$ for $\Delta_x^{-1} = 19$ when choosing the position in time of the source points according to [298]. This error was rather constant with respect to time. Of course, as the function under consideration decreases over time, the maximum absolute error also decreases. For even smaller Δ_x, e.g., $\Delta_x^{-1} = 21$, the error grows and is several times larger than the solution for which we are looking.

In case of poroelasticity, we tested the time-marching scheme for Examples 6.3 and 6.5. For the spatial parameters, we took $\Delta_x^{-1} \in [2,26] \cap \mathbb{N}$ with $\gamma = 1$ or $\gamma = 3$. Moreover, instead of $\delta_{\mathrm{TM}} = \sqrt{8}$, i.e., the diameter of $(-1,1)^2$, we also considered $\delta_{\mathrm{TM}} = \sqrt{2}$ and $\delta_{\mathrm{TM}} = 0.5$. Δ_t remained fixed as $\Delta_t = 0.01$. For the choice of source point locations, we compared the one given by Fig. 6.48 as well as another one, closer to Fig. 6.1. This means that the same Δ_x yields more points. The points at the boundary are chosen closer to the corners (minimal distance of $\frac{\Delta_x}{2}$ instead of Δ_x) and the points in Ω are chosen closer to the boundary (minimal distance of $\frac{\Delta_{x,0}}{2}$ instead of $\Delta_{x,0}$).

Comparing the results for different choices of source points, it turned out that source points according to $\gamma = 1$ yield better results than the ones for $\gamma = 3$. Similarly, we found that $\delta_{\mathrm{TM}} = \sqrt{2}$ performs better than the other two alternatives, but not in all cases. Moreover, a source point choice related to Fig. 6.1 usually performed better than a choice according to Fig. 6.48. However, the best results showed errors that were at least as large as 33 % of the analytical solution (pressure for $\Delta_x = \frac{1}{23}$). Even in that case, all other poroelastic quantities had errors which were several times larger than the analytical solution. We conclude that in order to apply this time-marching scheme, a much more sophisticated choice of source points is necessary, probably in a way that the set of source points is chosen adaptively to the problem under consideration. This corresponds to a comment by Prof. Dr. Fan in personal communication, in which he stated that the choice of suitable source points for parabolic differential equations is even more critical than for elliptic differential equations. For a (non-linear) algorithm allowing moving source points whose location is optimized, see [189].

Another possible explanation for the poor performance may be that, as we use some kind of time-marching here, a CFL-condition (see Remark 6.7) ought to be considered, which is not satisfied by our choices of Δ_x and Δ_t. This yields a further topic for future research.

6.6 Examples with Steeper Gradients

In all previous examples, the functions that we approximated were rather smooth with small gradients. In this section, we consider two examples with steeper gradients.

Example 6.8 Let $\Omega = (-1,1)^2$, $\Gamma = \partial\Omega$.

Approximate $u^{\mathrm{Si}}(x_1-1.05, x_2-1.05, t+0.05)$, $p^{\mathrm{Si}}(x_1-1.05, x_2-1.05, t+0.05)$ in $\Omega \times (0, t_{\mathrm{end}})$, given the initial condition $\zeta^0(x_1, x_2, 0) = \zeta^{\mathrm{Si}}(x_1-1.05, x_2-1.05, 1)$ on Ω and the mixed boundary conditions

$$
\begin{aligned}
&u^{\mathrm{BC}}(x_1, x_2, t) = u^{\mathrm{Si}}(x_1 - 1.05, x_2 - 1.05, t + 0.05)\ ,\\
&p^{\mathrm{BC}}(x_1, x_2, t) = p^{\mathrm{Si}}(x_1 - 1.05, x_2 - 1.05, t + 0.05)\ ,\\
&\text{for } x_1 \in \{-1, 1\}, x_2 \in (-1, 1), t \in (0, t_{\mathrm{end}})\ ,\\
&\boldsymbol{\sigma}^{\mathrm{pe,BC}}(x_1, x_2, t)n(x) = \boldsymbol{\sigma}^{\mathrm{pe,Si}}(x_1 - 1.05, x_2 - 1.05, t + 0.05)\begin{pmatrix} 0 \\ x_2 \end{pmatrix}\ ,\\
&v_{f,2}^{\mathrm{BC}}(x_1, x_2, t) = x_2 \partial_{x_2} p^{\mathrm{Si}}(x_1 - 1.05, x_2 - 1.05, t + 0.05)\ ,\\
&\text{for } x_1 \in (-1, 1), x_2 \in \{-1, 1\}, t \in (0, t_{\mathrm{end}})\ .
\end{aligned}
$$

Example 6.9 Let $\Omega = (-1,1)^2$, $\Gamma = \partial\Omega$.

Approximate $u^{\mathrm{fi}}(x_1-1.05, x_2-1.05, t+0.05)$, $p^{\mathrm{fi}}(x_1-1.05, x_2-1.05, t+0.05)$ in $\Omega \times (0, t_{\mathrm{end}})$, given the initial condition $\zeta^0(x_1, x_2, 0) = \zeta^{\mathrm{fi}}(x_1-1.05, x_2-1.05, 0+1)$ on Ω and the mixed boundary conditions

$$
\begin{aligned}
&u^{\mathrm{BC}}(x_1, x_2, t) = u^{\mathrm{fi}}(x_1 - 1.05, x_2 - 1.05, t + 0.05)\ ,\\
&p^{\mathrm{BC}}(x_1, x_2, t) = p^{\mathrm{fi}}(x_1 - 1.05, x_2 - 1.05, t + 0.05)\ ,\\
&\text{for } x_1 \in \{-1, 1\}, x_2 \in (-1, 1), t \in (0, t_{\mathrm{end}})\ ,\\
&\boldsymbol{\sigma}^{\mathrm{pe,BC}}(x_1, x_2, t)n(x) = \boldsymbol{\sigma}^{\mathrm{pe,fi}}(x_1 - 1.05, x_2 - 1.05, t + 0.05)\begin{pmatrix} 0 \\ x_2 \end{pmatrix}\ ,\\
&v_{f,2}^{\mathrm{BC}}(x_1, x_2, t) = x_2 \partial_{x_2} p^{\mathrm{fi}}(x_1 - 1.05, x_2 - 1.05, t + 0.05)\ ,\\
&\text{for } x_1 \in (-1, 1), x_2 \in \{-1, 1\}, t \in (0, t_{\mathrm{end}})\ .
\end{aligned}
$$

The difference between Examples 6.3 and 6.5 on the one hand and Examples 6.8 and 6.9 is solely in where the singularity of the fundamental solution is placed. In the latter two examples, the singularity is located close to the boundary Γ and close to $t = 0$, resulting in steep gradients with respect to space as well as time, as can be seen in Figs. 6.49–6.51 and 6.54–6.56 respectively. The directions of the vector-valued quantities u and v_f are given in Figs. 6.52 and 6.53 as well as 6.57 and 6.58.

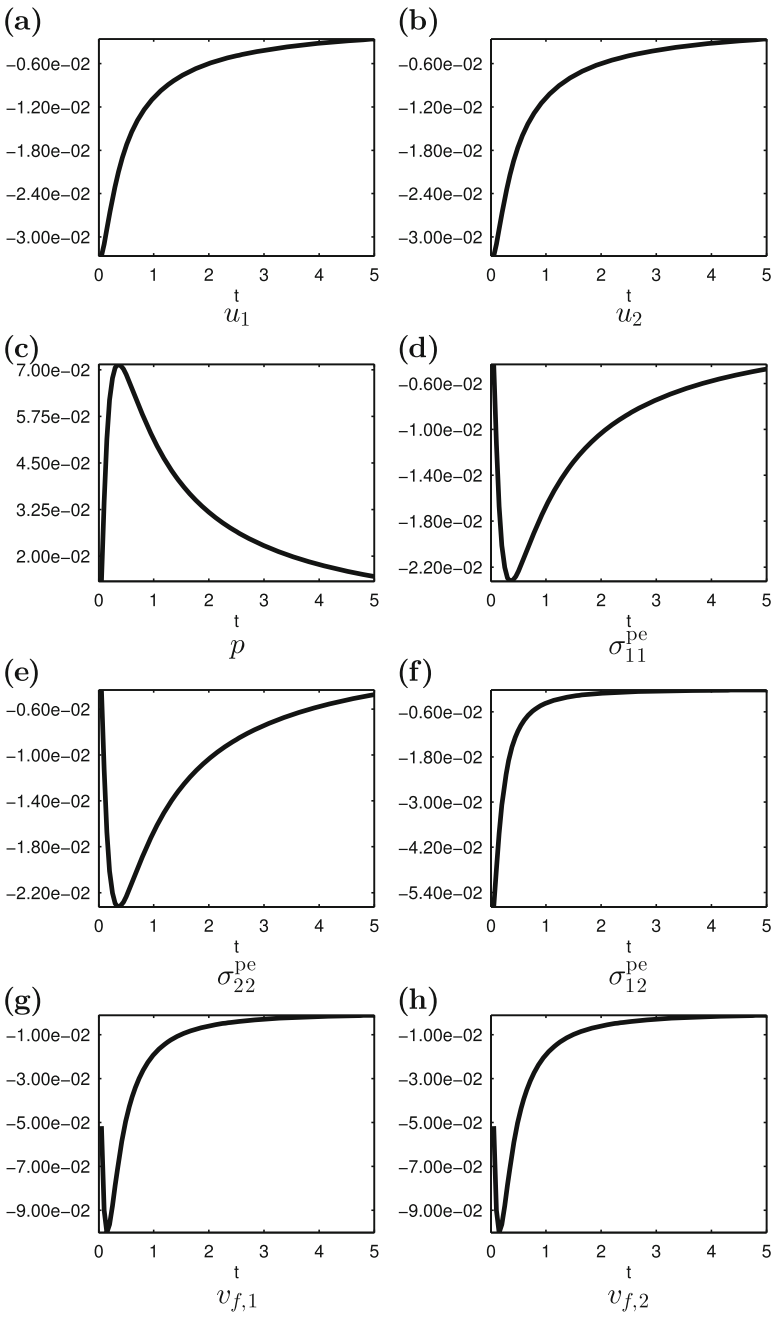

Fig. 6.49 Time development of the poroelastic quantities for Example 6.8 evaluated at the origin $x = 0$. Please note that different scales are used for different quantities as indicated by the different scales on the ordinates. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

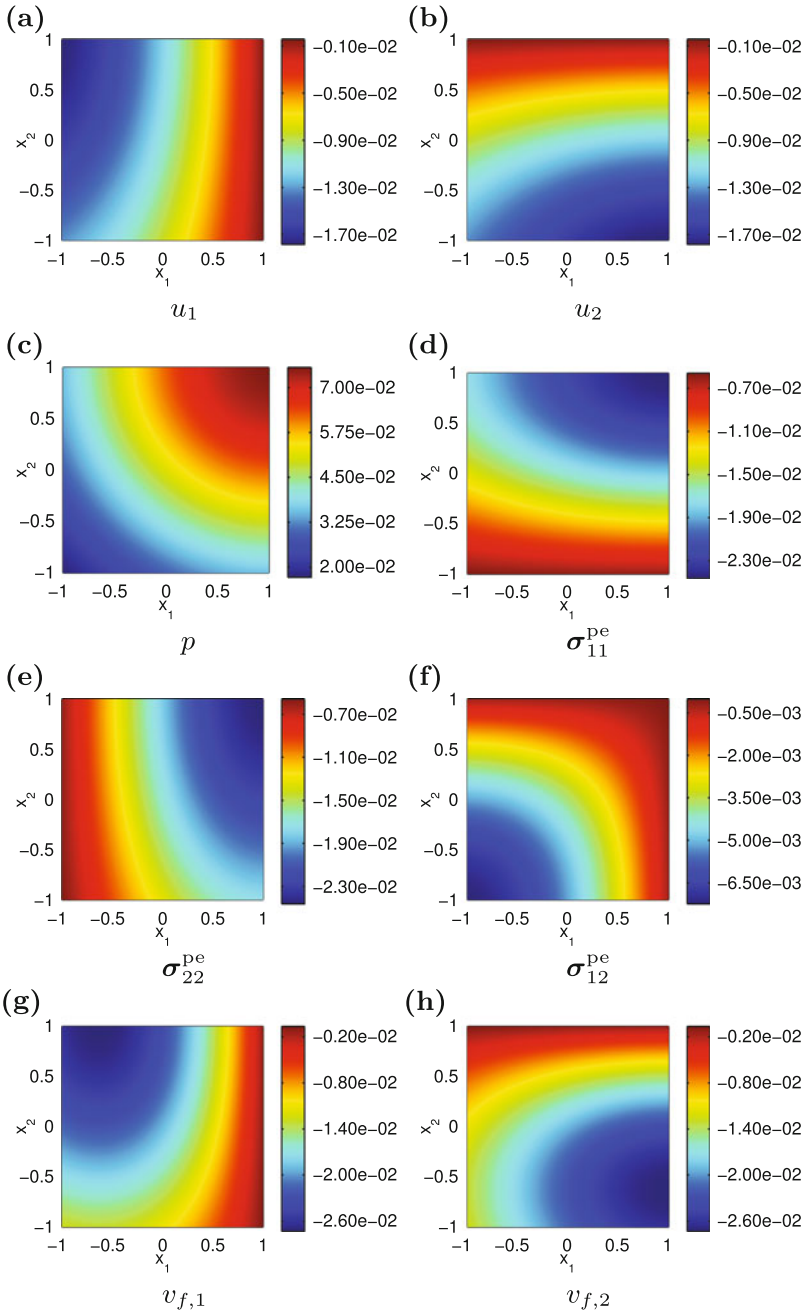

Fig. 6.50 Pseudocolor plots of the poroelastic quantities for Example 6.8 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Please note that different scales are used for different quantities as indicated by the different scales on the colorbars. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

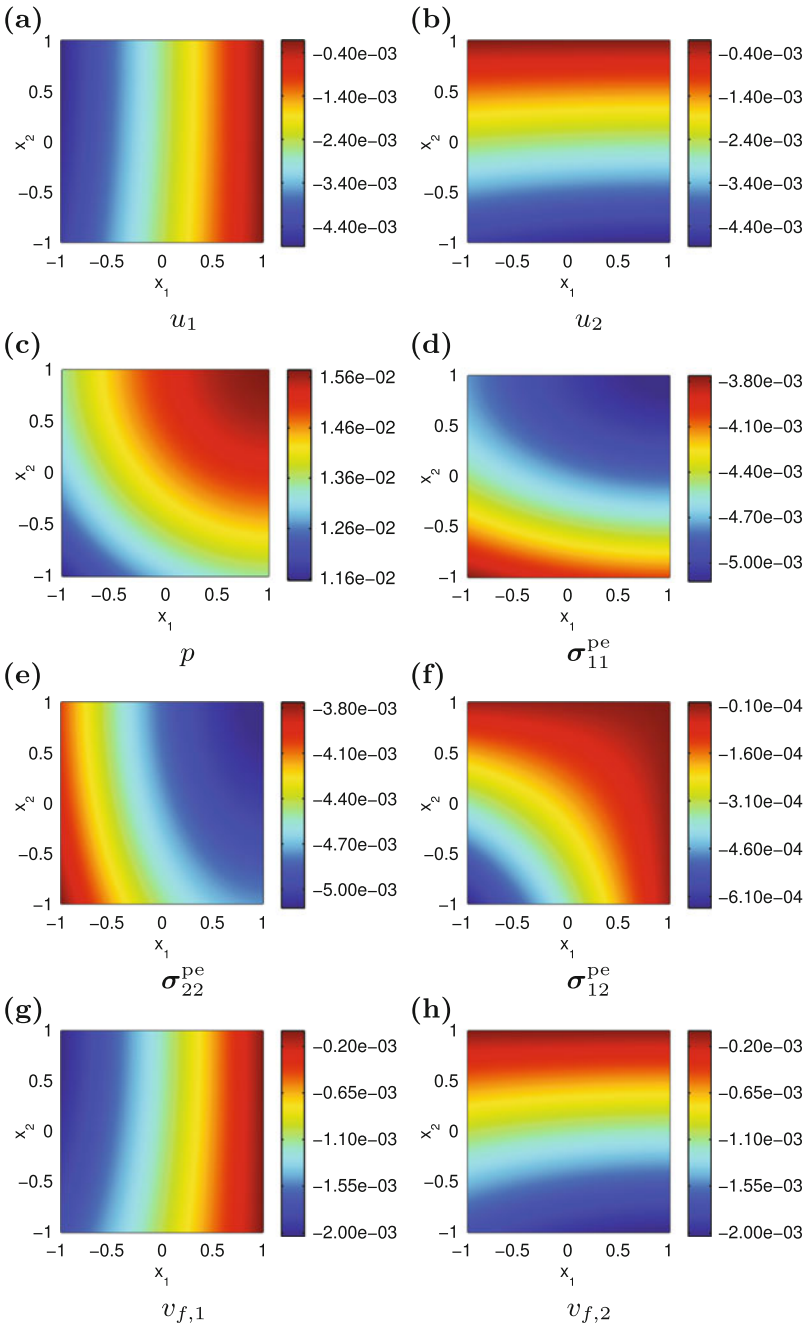

Fig. 6.51 Pseudocolor plots of the poroelastic quantities for Example 6.8 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Please note that different scales are used for different quantities as indicated by the different scales on the colorbars. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

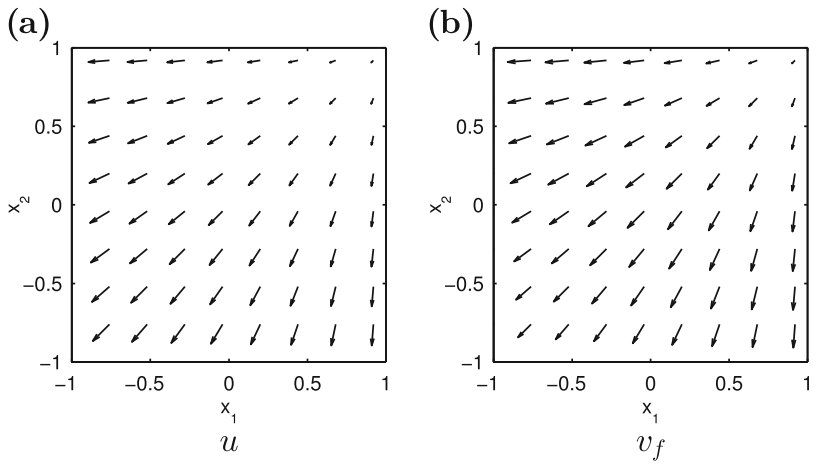

Fig. 6.52 Directions of the vector-valued poroelastic quantities u and v_f for Example 6.8 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Please compare Fig. 6.50a, b and g, h as the above plots do not give sufficient information about the strength of the fields

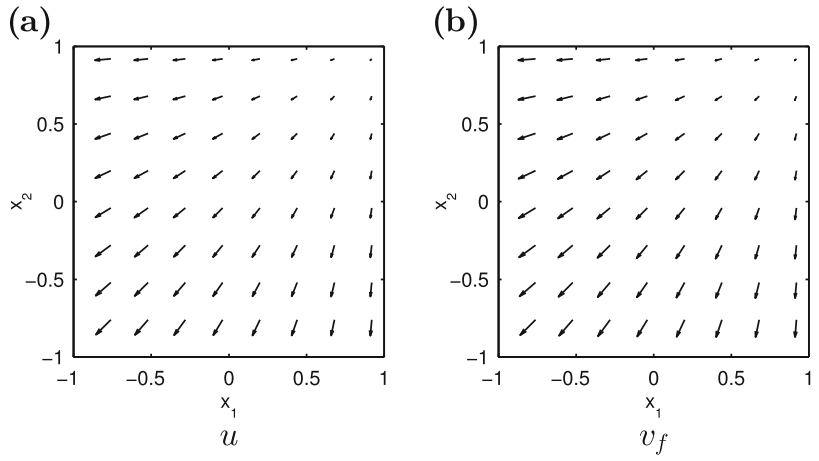

Fig. 6.53 Directions of the vector-valued poroelastic quantities u and v_f for Example 6.8 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Please compare Fig. 6.51a, b and g, h as the above plots do not give sufficient information about the strength of the fields

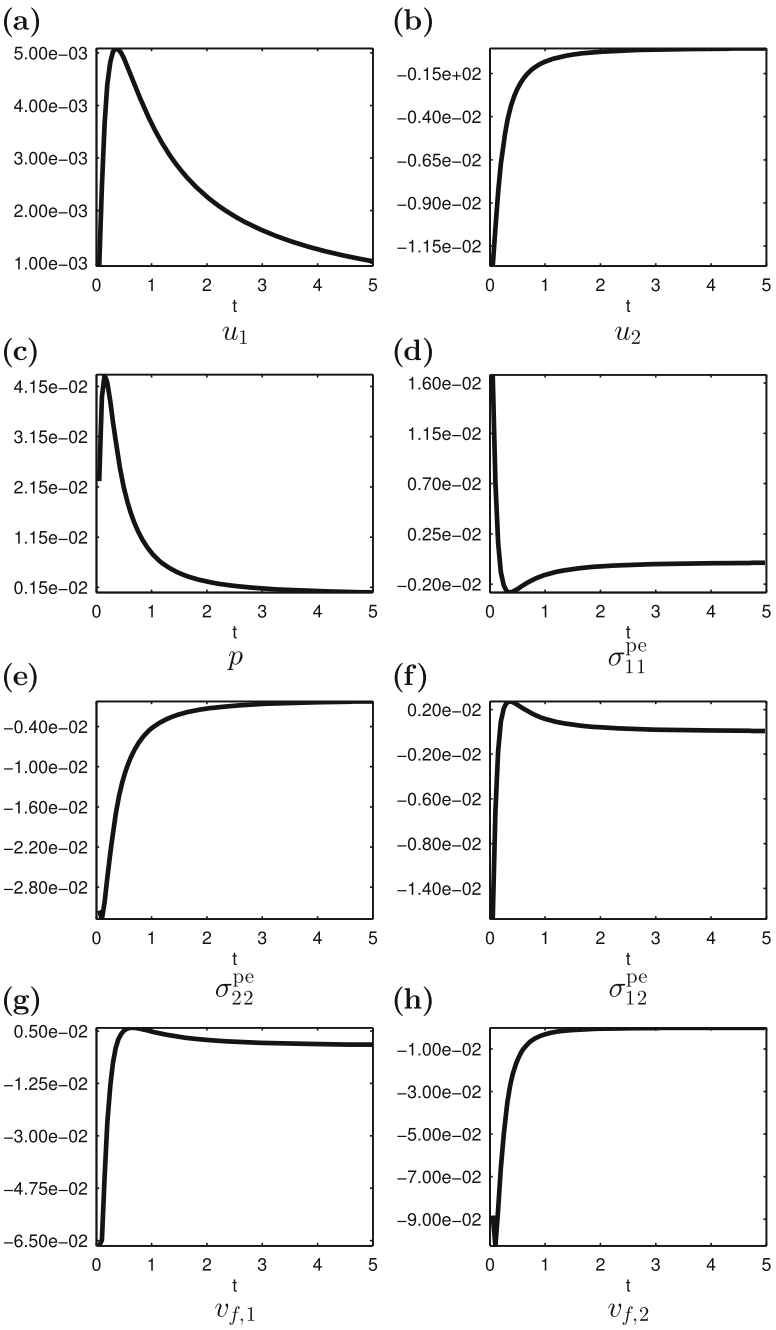

Fig. 6.54 Time development of the poroelastic quantities for Example 6.9 evaluated at the origin $x = 0$. Please note that different scales with different signs are used for different quantities as indicated by the different scales on the ordinates. All quantities are dimensionless

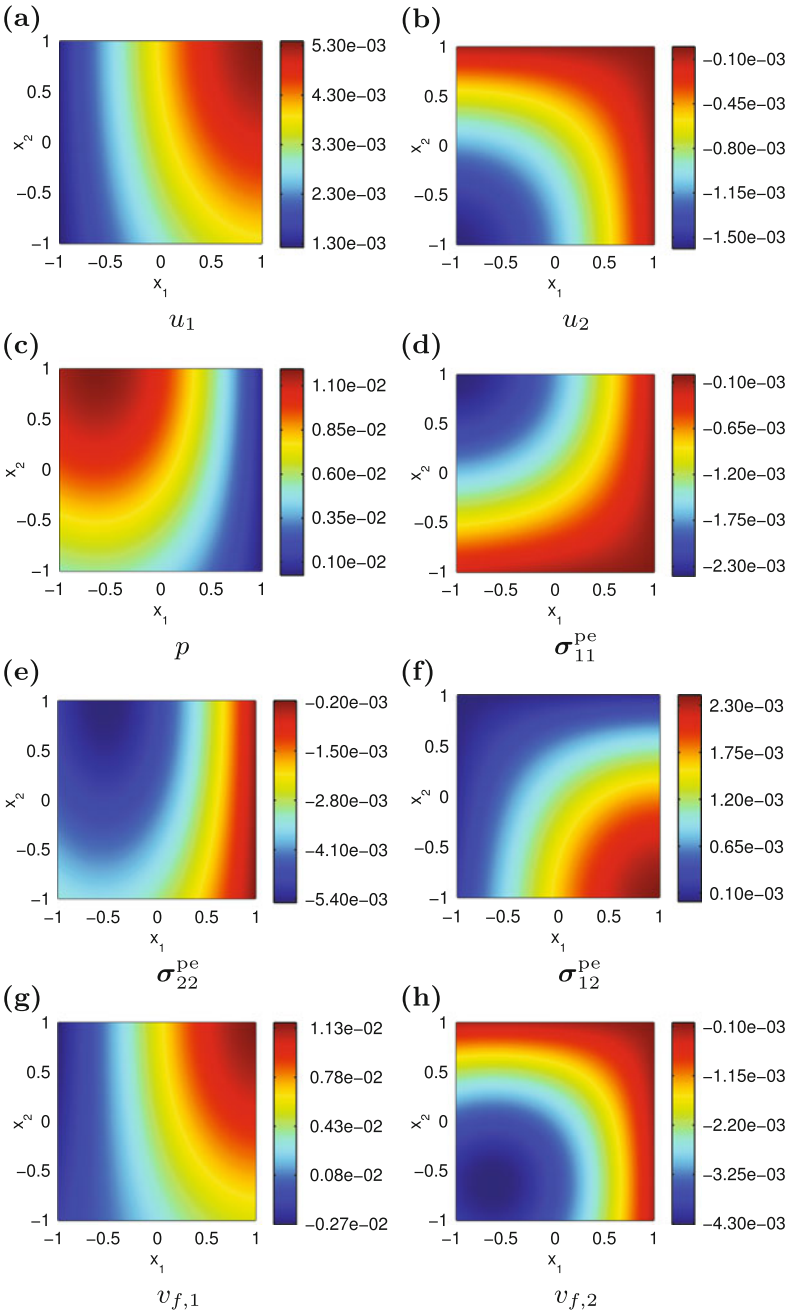

Fig. 6.55 Pseudocolor plots of the poroelastic quantities for Example 6.9 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Please note that different scales with different signs are used for different quantities as indicated by the different scales on the colorbars. All quantities are dimensionless

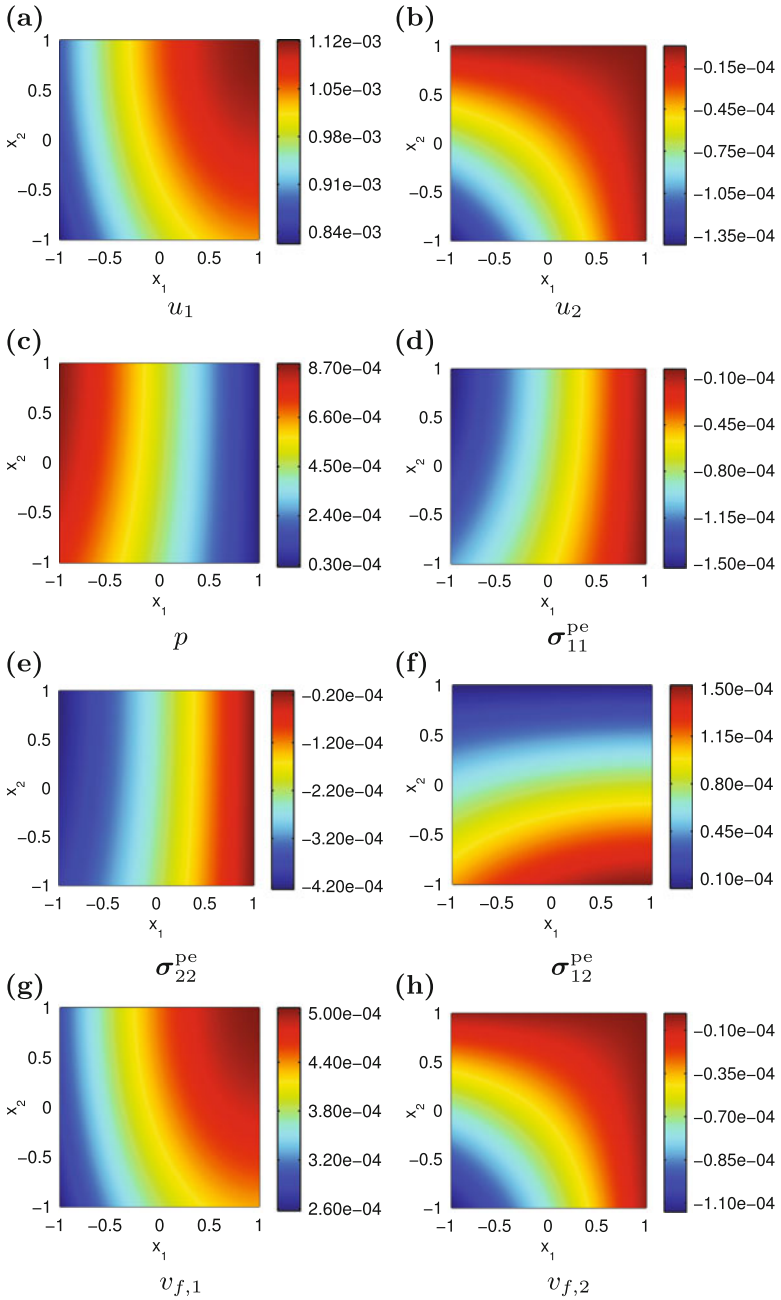

Fig. 6.56 Pseudocolor plots of the poroelastic quantities for Example 6.9 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Please note that different scales with different signs are used for different quantities as indicated by the different scales on the colorbars. All quantities are dimensionless

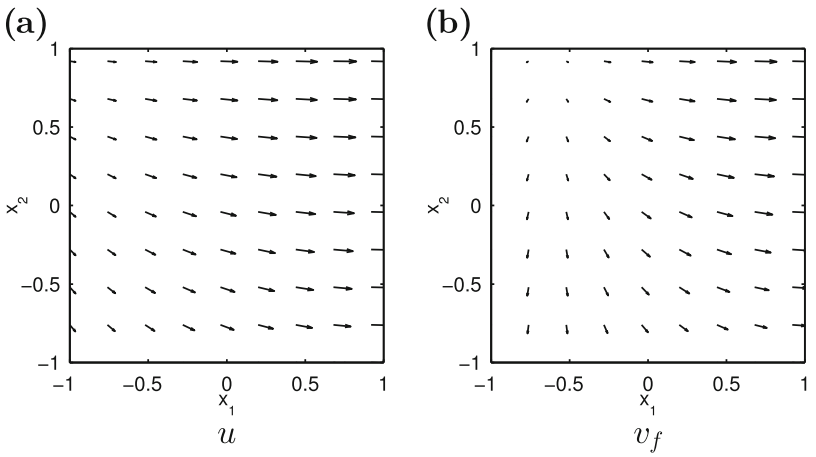

Fig. 6.57 Directions of the vector-valued poroelastic quantities u and v_f for Example 6.9 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 1$. Please compare Fig. 6.55a, b and g, h as the above plots do not give sufficient information about the strength of the fields

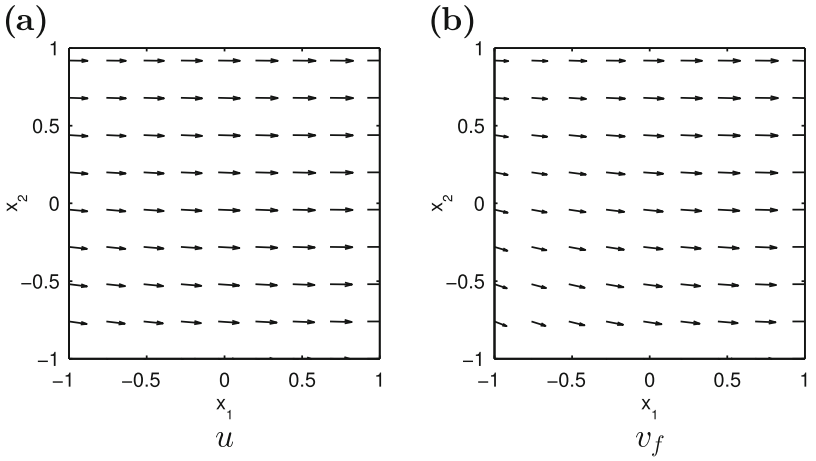

Fig. 6.58 Directions of the vector-valued poroelastic quantities u and v_f for Example 6.9 evaluated on the square $\Omega = (-1, 1)^2$ at time $t = 5$. Please compare Fig. 6.56a, b and g, h as the above plots do not give sufficient information about the strength of the fields

Table 6.8 Errors in the poroelastic quantities for Example 6.8. Parameters are $\Delta_x = \frac{1}{3}$ ($I = 6$), $\Delta_y = \frac{1}{2}$ ($M = 4$), $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}$ with respect to the domain as well as the maximum relative error $\mathrm{E}^{\mathrm{rel}}_{\mathrm{max}}$ as defined in (6.15), (6.16), and (6.10b), and points of $\mathfrak{I}_2$ as well as times of $\mathfrak{I}_t$ at which they are obtained

Quantity	$\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}$	x_1	x_2	$\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}$	t	$\mathrm{E}^{\mathrm{rel}}_{\mathrm{max}}$	x_1	x_2	t
u_1	7.4890e–02	1.0	−1.0	1.6976e–03	5.0	1.6899e–01	1.0	−1.0	3.95
u_2	2.3636e–01	0.0	1.0	4.1571e–03	5.0	5.3083e–01	−0.5	1.0	0.05
p	4.7150e–02	−0.5	−1.0	1.1937e–03	1.0	1.5010e+04	−1.0	−1.0	0.05
$\boldsymbol{\sigma}^{\mathrm{pe}}_{11}$	9.6149e–02	−1.0	−1.0	5.3294e–03	5.0	6.9675e+04	−1.0	−1.0	0.05
$\boldsymbol{\sigma}^{\mathrm{pe}}_{22}$	1.1289e–01	−1.0	−1.0	4.1346e–03	5.0	9.5238e+04	−1.0	−1.0	0.05
$\boldsymbol{\sigma}^{\mathrm{pe}}_{12}$	1.6945e+00	1.0	1.0	7.2644e–02	5.0	1.5149e+02	1.0	1.0	5.00
$v_{f,1}$	5.6466e+00	1.0	−1.0	4.4662e–02	5.0	3.6874e+04	−1.0	−1.0	0.05
$v_{f,2}$	3.5054e+00	−1.0	1.0	2.0780e–02	5.0	2.1396e+04	−1.0	−1.0	0.05

Table 6.9 Errors in the poroelastic quantities for Example 6.8. Parameters are $\Delta_x = \frac{1}{5}$ ($I = 10$), $\Delta_y = \frac{1}{3}$ ($M = 6$), $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}$ with respect to the domain as well as the maximum relative error $\mathrm{E}^{\mathrm{rel}}_{\mathrm{max}}$ as defined in (6.15), (6.16), and (6.10b), and points of $\mathfrak{I}_2$ as well as times of $\mathfrak{I}_t$ at which they are obtained

Quantity	$\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,t}}$	x_1	x_2	$\mathrm{E}^{\mathrm{rel}}_{\mathrm{rms,x}}$	t	$\mathrm{E}^{\mathrm{rel}}_{\mathrm{max}}$	x_1	x_2	t
u_1	1.0288e–02	1.0	−1.0	4.8377e–04	5.0	3.5955e–02	0.5	−1.0	0.05
u_2	1.5147e–01	0.0	1.0	8.0351e–04	1.0	2.8176e–01	0.0	1.0	0.05
p	5.3901e–02	0.0	−1.0	1.3138e–03	1.0	2.6417e+02	−0.5	−1.0	0.05
$\boldsymbol{\sigma}^{\mathrm{pe}}_{11}$	7.9566e–02	−1.0	−1.0	2.7659e–03	5.0	4.1038e+04	−1.0	−1.0	0.05
$\boldsymbol{\sigma}^{\mathrm{pe}}_{22}$	1.3469e–02	−1.0	−1.0	1.0353e–03	5.0	3.4882e+03	−1.0	−1.0	0.05
$\boldsymbol{\sigma}^{\mathrm{pe}}_{12}$	1.9955e–02	1.0	1.0	1.2255e–02	5.0	2.3657e+01	1.0	1.0	4.95
$v_{f,1}$	1.5057e+00	1.0	−0.5	6.2754e–02	5.0	2.9093e+03	−1.0	−1.0	0.05
$v_{f,2}$	2.5438e–01	−1.0	1.0	5.2639e–02	5.0	9.1127e+02	−1.0	−1.0	0.05

In order to approximate the solutions to Examples 6.8 and 6.9, we fixed $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$ and considered the spatial parameter combinations $\Delta_x = \frac{1}{3}$, $\Delta_y = \frac{1}{2}$, $\gamma = 1$, and $\Delta_x = \frac{1}{5}$, $\Delta_y = \frac{1}{3}$, $\gamma = 1$. The first one is the same as for most other considerations in this chapter in which spatial parameters are not varied. The second one showed the best performance of all tested combinations of spatial parameters with regard to Examples 6.2–6.5. The approximation errors for those parameter combinations and both examples with steep gradients are given in Tables 6.8–6.11.

As we can see in Tables 6.8–6.11, the errors are several orders of magnitude larger, compared to the results for Examples 6.2–6.5. Even for the better performing parameters $\Delta_x = \frac{1}{5}$, $\Delta_y = \frac{1}{3}$, we still have rooted mean square errors of at least 1% (u_1 in Example 6.8, see Table 6.9) and up to about 900 % ($\boldsymbol{\sigma}^{\mathrm{pe}}_{11}$ in Example 6.9,

Table 6.10 Errors in the poroelastic quantities for Example 6.9. Parameters are $\Delta_x = \frac{1}{3}$ ($I = 6$), $\Delta_y = \frac{1}{2}$ ($M = 4$), $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as well as the maximum relative error $\mathrm{E}^{\text{rel}}_{\text{max}}$ as defined in (6.15), (6.16), and (6.10b), and points of $\mathfrak{I}_2$ as well as times of $\mathfrak{I}_t$ at which they are obtained

Quantity	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}$	x_1	x_2	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t	$\mathrm{E}^{\text{rel}}_{\text{max}}$	x_1	x_2	t
u_1	5.1968e–02	−1.0	1.0	2.3061e–02	4.0	2.1448e+04	−1.0	−1.0	0.05
u_2	3.2979e+00	1.0	1.0	3.0525e–01	4.0	2.2936e+02	1.0	1.0	5.00
p	1.8362e+00	1.0	−1.0	6.2484e–01	5.0	1.1059e+03	−0.5	−1.0	0.05
$\boldsymbol{\sigma}^{\text{pe}}_{11}$	8.8587e+00	1.0	−1.0	8.4347e–01	5.0	1.0300e+03	1.0	−1.0	0.90
$\boldsymbol{\sigma}^{\text{pe}}_{22}$	3.4467e+00	1.0	−1.0	2.4163e–01	5.0	2.4407e+01	1.0	−1.0	4.95
$\boldsymbol{\sigma}^{\text{pe}}_{12}$	5.7665e+01	−1.0	1.0	4.9368e–01	5.0	6.3839e+01	1.0	1.0	4.95
$v_{f,1}$	1.9018e+01	−0.5	−1.0	2.2503e–01	5.0	1.8163e+03	−1.0	−1.0	0.05
$v_{f,2}$	1.1100e+01	−1.0	1.0	1.3748e+00	5.0	1.8713e+03	−1.0	−1.0	0.05

Table 6.11 Errors in the poroelastic quantities for Example 6.9. Parameters are $\Delta_x = \frac{1}{5}$ ($I = 10$), $\Delta_y = \frac{1}{3}$ ($M = 6$), $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 5$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as well as the maximum relative error $\mathrm{E}^{\text{rel}}_{\text{max}}$ as defined in (6.15), (6.16), and (6.10b), and points of $\mathfrak{I}_2$ as well as times of $\mathfrak{I}_t$ at which they are obtained

Quantity	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}$	x_1	x_2	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t	$\mathrm{E}^{\text{rel}}_{\text{max}}$	x_1	x_2	t
u_1	8.4166e–02	−0.5	−1.0	1.3914e–02	4.0	8.7664e+03	−1.0	−1.0	0.05
u_2	4.4837e–01	0.5	1.0	1.4773e–01	5.0	7.2598e+01	1.0	1.0	4.90
p	9.7536e–01	0.5	−1.0	2.3221e–01	5.0	1.2509e+03	−0.5	−1.0	0.05
$\boldsymbol{\sigma}^{\text{pe}}_{11}$	9.1796e+00	1.0	−1.0	9.4946e–01	5.0	5.7807e+02	1.0	−1.0	0.90
$\boldsymbol{\sigma}^{\text{pe}}_{22}$	4.0378e–01	1.0	−1.0	1.7704e–01	5.0	1.3215e+01	1.0	1.0	5.00
$\boldsymbol{\sigma}^{\text{pe}}_{12}$	3.1073e–01	−1.0	1.0	2.4872e–01	5.0	1.7941e+01	−1.0	1.0	4.95
$v_{f,1}$	2.9498e+00	−1.0	0.0	2.4556e+00	5.0	1.1697e+03	0.5	−1.0	0.05
$v_{f,2}$	7.7834e–01	−1.0	1.0	1.0177e+01	5.0	9.3281e+02	1.0	1.0	4.95

see Table 6.11). The maximum relative errors are even worse, reaching values of almost 10^5.

These results are not unexpected. Indeed, it is well known that solutions with steep gradients or even discontinuities yield difficulties in numerical modeling, e.g. strong oscillations, overshoots, or undershoots, and often require some kind of stabilization (see [11, 16, 171, 270] and the references therein). In the context of MFSs, it is often necessary to adapt the choice of source points to handle such problems (see [24] for an example of an adaptive method based on conformal mapping methods for Helmholtz equations). Developing such an adaptive method is out of the scope of this thesis, but the results discussed in this section motivate further research on this topic.

Remark 6.10 A commonly used example to test numerical solution schemes for the QEP is Mandel's problem [2, 61]. Mandel [185] assumed that the following initial and boundary conditions are valid:

$$\zeta(x_1, x_2, 0) = 0, \qquad x \in (-1, 1)^2 , \tag{6.23a}$$

$$\begin{aligned} \sigma_{11}^{\mathrm{pe}}(x_1, x_2, t) &= \sigma_{12}^{\mathrm{pe}}(x_1, x_2, t) \\ &= p(x_1, x_2, t) = 0, \quad x_1 \in \{-1, 1\}, x_2 \in (-1, 1) , \end{aligned} \tag{6.23b}$$

$$\sigma_{12}^{\mathrm{pe}}(x_1, x_2, t) = v_{f,2}(x_1, x_2, t) = 0, \qquad x_1 \in (-1, 1), x_2 \in \{-1, 1\} , \tag{6.23c}$$

$$u_2(x_1, x_2, t) = u_2(t), \qquad x_1 \in (-1, 1), x_2 \in \{-1, 1\} , \tag{6.23d}$$

$$\int_{-1}^{1} \sigma_{22}^{\mathrm{pe}}(x_1, x_2, t)\, dx_1 = -2F\, H(t), \qquad x_2 \in \{-1, 1\} , \tag{6.23e}$$

with a constant $F \in \mathbb{R}^+$, representing a given force which is applied to the domain.

Mandel's problem provides an analytical solution [2, 61], which, for small times, has steep gradients at the boundary of the domain [221] and also steep gradients with respect to time near $t = 0$. Due to these two features, especially the rapid growth with respect to time, finding a solution to Mandel's problem with an MFS turns out to be difficult. These difficulties may be conquered by a very careful choice of source points. It may even be necessary to implement a non-linear algorithm that searches for the best source point locations [189]. Therefore, investigating Mandel's problem is out of the scope of this thesis.

Another related problem on a half-ball is Cryer's problem [168]. Both of these examples demonstrate a non-monotonic pressure response which is typical for poroelastic models and absent in simple diffusion problems. This behavior is known as Mandel-Cryer effect or Noordbergum effect (see [2] and the references therein). It is also present in the examples considered within this chapter.

6.7 An Example on the Cube $(-1, 1)^3$

In order to demonstrate that the MFS in quasistatic poroelasticity is also applicable on a three-dimensional domain, we consider the following example:

Example 6.11 Let $\Omega = (-1, 1)^3$, $\Gamma = \partial\Omega$.

Approximate $u^{\mathrm{Si}}(x_1 - 2, x_2 - 2, x_3 - 2, t + 1)$, $p^{\mathrm{Si}}(x_1 - 2, x_2 - 2, x_3 - 2, t + 1)$ in $\Omega \times (0, t_{\mathrm{end}})$, given the initial condition $\zeta^0(x_1, x_2, x_3, 0) = \zeta^{\mathrm{Si}}(x_1 - 2, x_2 - 2, x_3 - 2, 1)$

on Ω and the mixed boundary conditions

$$
\begin{aligned}
&u^{\mathrm{BC}}(x_1,x_2,x_3,t) = u^{\mathrm{Si}}(x_1-2,x_2-2,x_3-2,t+1)\ ,\\
&p^{\mathrm{BC}}(x_1,x_2,x_3,t) = p^{\mathrm{Si}}(x_1-2,x_2-2,x_3-2,t+1)\ ,\\
&\text{for } x \in (\{-1,1\}\times(-1,1)\times(-1,1)) \cup ((-1,1)\times\{-1,1\}\times(-1,1))\ ,\\
&\qquad t \in (0,t_{\mathrm{end}})\ ,\\
&\boldsymbol{\sigma}^{\mathrm{pe,BC}}(x_1,x_2,x_3,t)n(x) = \boldsymbol{\sigma}^{\mathrm{pe,Si}}(x_1-2,x_2-2,x_3-2,t+1)\begin{pmatrix}0\\0\\x_3\end{pmatrix}\ ,\\
&v_{f,3}^{\mathrm{BC}}(x_1,x_2,x_3,t) = x_3\partial_{x_3}p^{\mathrm{Si}}(x_1-2,x_2-2,x_3-2,t+1)\ ,\\
&\text{for } x \in (-1,1)\times(-1,1)\times\{-1,1\}, t\in(0,t_{\mathrm{end}})\ .
\end{aligned}
$$

Please note that the cartesian product of two sets yields tuple, such that the above given cartesian products give different faces of the cube.

The development in time for u and v_f evaluated at the origin $x = (0,0)$ is given in Fig. 6.59. As we can see, all components of u are given by the same function, a statement that is also true for the components of v_f, as can be expected. However, the scales for u and v_f differ. All of these quantities show an almost linear growth with increasing time.

Figure 6.60 depicts the behavior of p and the components of $\boldsymbol{\sigma}^{\mathrm{pe}}$ in time. We recognize that the diagonal elements of $\boldsymbol{\sigma}^{\mathrm{pe}}$ behave qualitatively much like p, showing a maximum at about $t = 0.5$, but differ quantitatively. The non-diagonal elements of $\boldsymbol{\sigma}^{\mathrm{pe}}$, which share the same behavior among themselves, increase monotonically and have a different sign as well as a smaller absolute value.

Figures 6.61 and 6.63 show u and v_f for fixed time $t = 0.2$ and $t = 1$, respectively, for two planar intersections through Ω, given by $x_1 = 0$ and $x_3 = 0$. The same scale is used for all components of u and v_f, respectively. The symmetry of the problem is clearly visible by comparing the first and third components of each of those vector-valued quantities. The second components seem not to share this symmetry. However, the shown figure is misleading, at it does not show the intersection with the plane $x_2 = 0$. Showing this intersection would reveal the missing symmetry, but would make it hard to distinguish between the intersections. Therefore, we decided to only depict two intersections.

Figures 6.62 and 6.64 show p and the components of $\boldsymbol{\sigma}^{\mathrm{pe}}$ for fixed time $t = 0.2$ and $t = 1$ in the same planar intersections as before. Again, different scales are used here, with the ones used for $\boldsymbol{\sigma}_{11}^{\mathrm{pe}}$, $\boldsymbol{\sigma}_{22}^{\mathrm{pe}}$, and $\boldsymbol{\sigma}_{33}^{\mathrm{pe}}$ being the same. The same is true for the scale used in the depiction of $\boldsymbol{\sigma}_{12}^{\mathrm{pe}}$, $\boldsymbol{\sigma}_{13}^{\mathrm{pe}}$, and $\boldsymbol{\sigma}_{23}^{\mathrm{pe}}$. Once more, we recognize the symmetry of the problem and the seemingly missing symmetry can be explained as before.

Collocation points and source points for $(-1,1)^3$ are chosen analogously to the choice on $(-1,1)^2$. On each face of the cube, collocation points are chosen in a pattern similar to the choice of collocation points for the initial condition in Fig. 6.1.

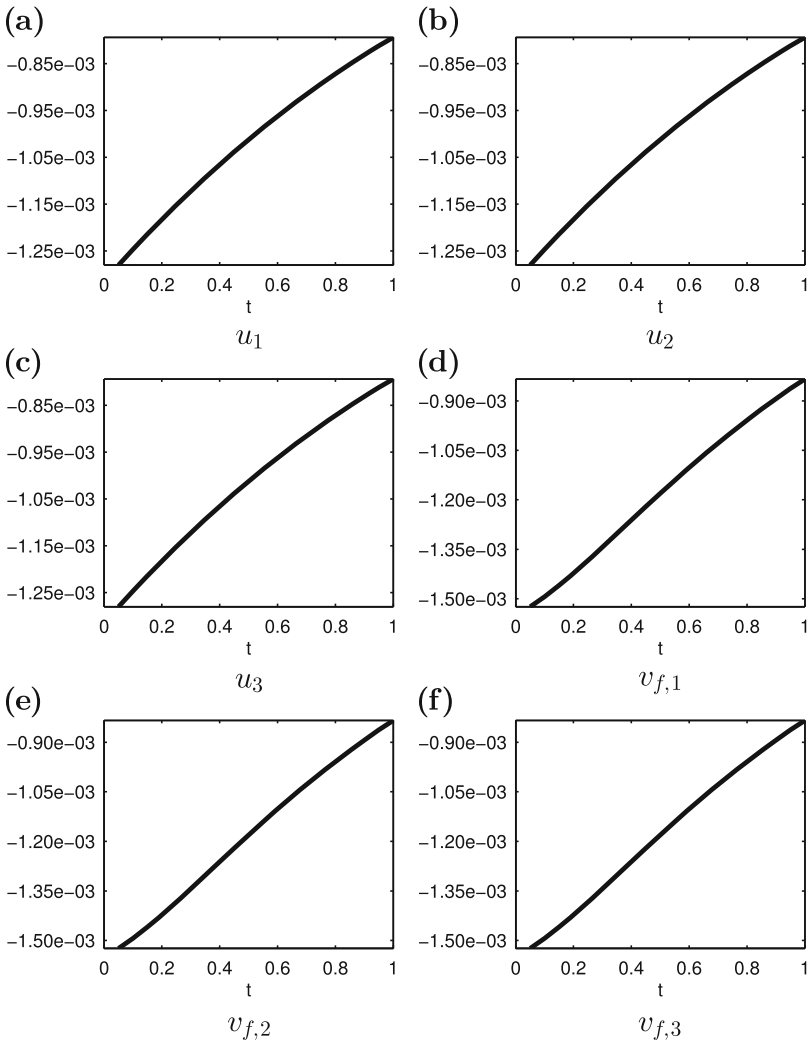

Fig. 6.59 Time development of u and v_f for Example 6.11 evaluated at the origin $x = 0$. Please note that different scales are used for different quantities as indicated by the different scales on the ordinates. All quantities are dimensionless

For example, for the face at $x_3 = 1$, we have the collocation points

$$\left\{x_1 : x_1 = -1 + \left(i - \frac{1}{2}\right)\Delta_x, i \in \mathbb{N}, i \leq I_{\text{int}}\right\}$$
$$\times \left\{x_2 : x_2 = -1 + \left(j - \frac{1}{2}\right)\Delta_x, j \in \mathbb{N}, j \leq I_{\text{int}}\right\}$$
$$\times \{x_3 = 1\}. \tag{6.24}$$

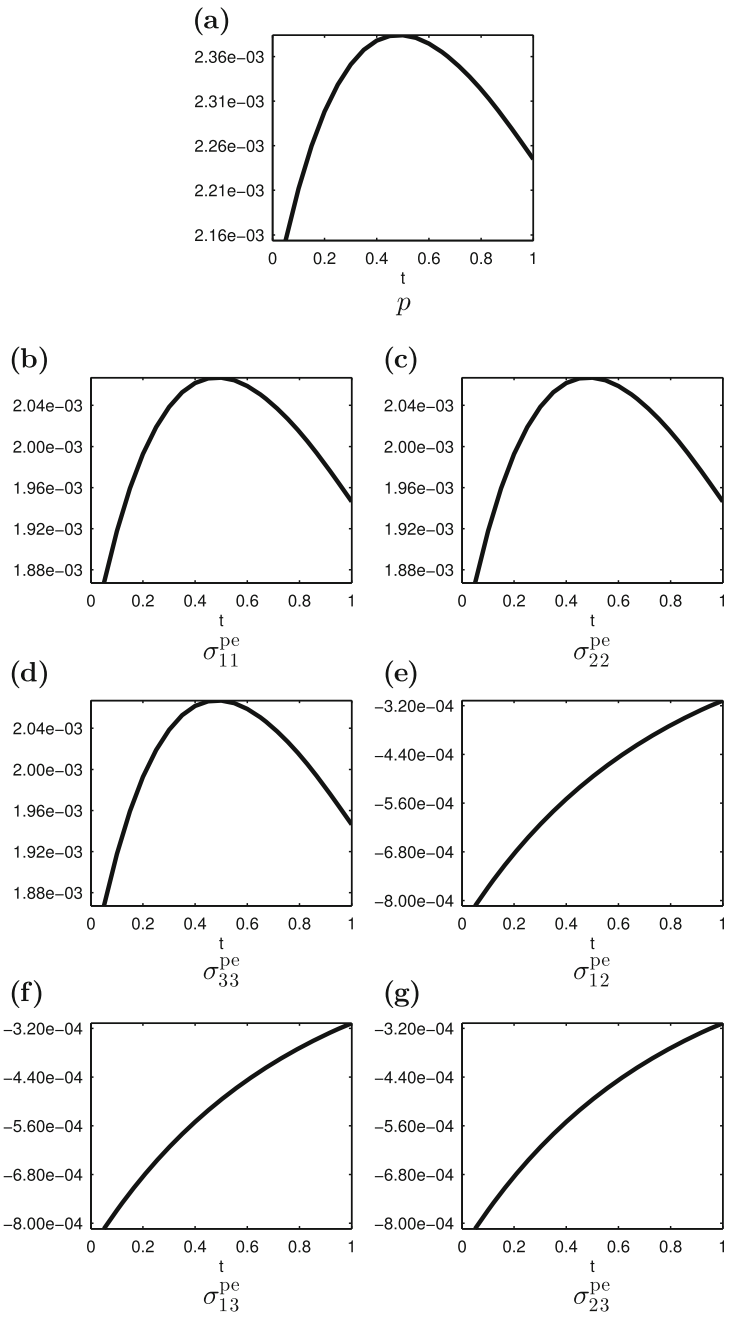

Fig. 6.60 Time development of p and $\boldsymbol{\sigma}^{\mathrm{pe}}$ for Example 6.11 evaluated at the origin $x = 0$. Please note that different scales are used for different quantities as indicated by the different scales on the ordinates. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

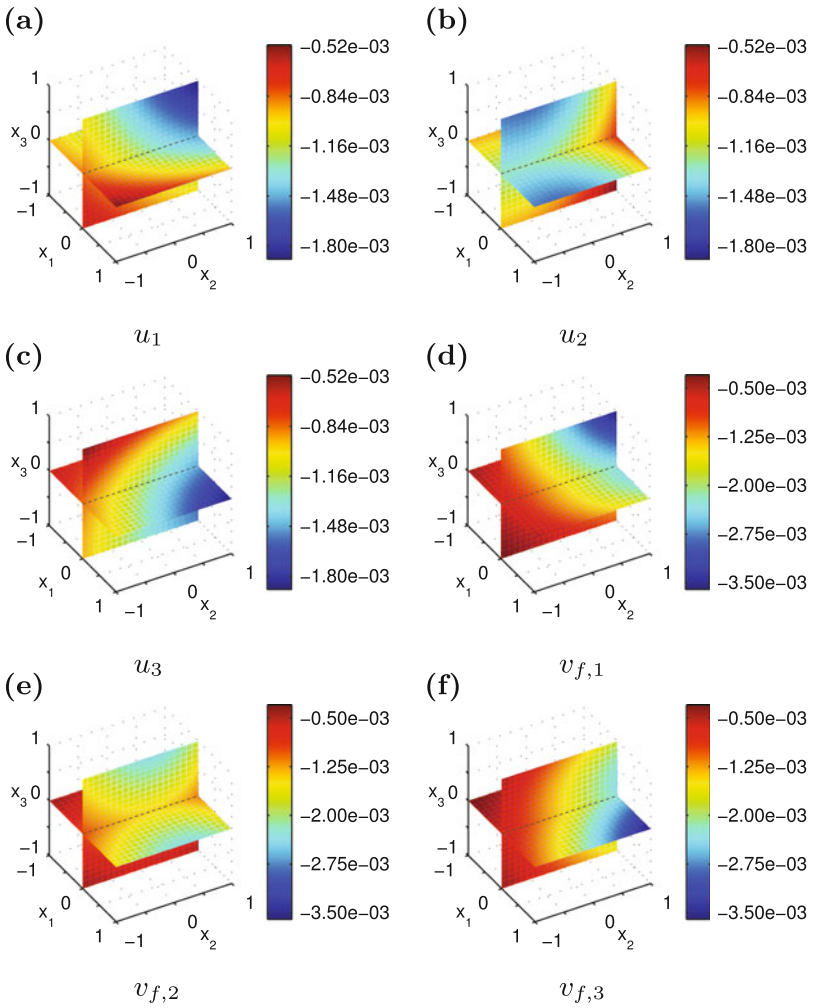

Fig. 6.61 Colored surface plots of u and v_f for Example 6.11 evaluated on the cube $\Omega = (-1, 1)^3$ at time $t = 1$. Please note that different scales are used for different quantities as indicated by the different scales on the colorbars. All quantities are dimensionless

For the initial condition, the definition for $\Delta_{x,0}$ is adjusted to a three-dimensional setting under the same conditions as before, i.e., having approximately as much collocation points in $\Omega \times \{t = 0\}$ as in $\Gamma \times (0, t_{\text{end}})$. Consequently, we define for a three-dimensional setting $I_{\text{int}} = \left\lfloor \sqrt[3]{6 I^2 J} \right\rfloor$. This yields a three-dimensional cartesian grid in the whole cube such that the points closest to the faces have a distance of $0.5\Delta_{x,0}$ to them.

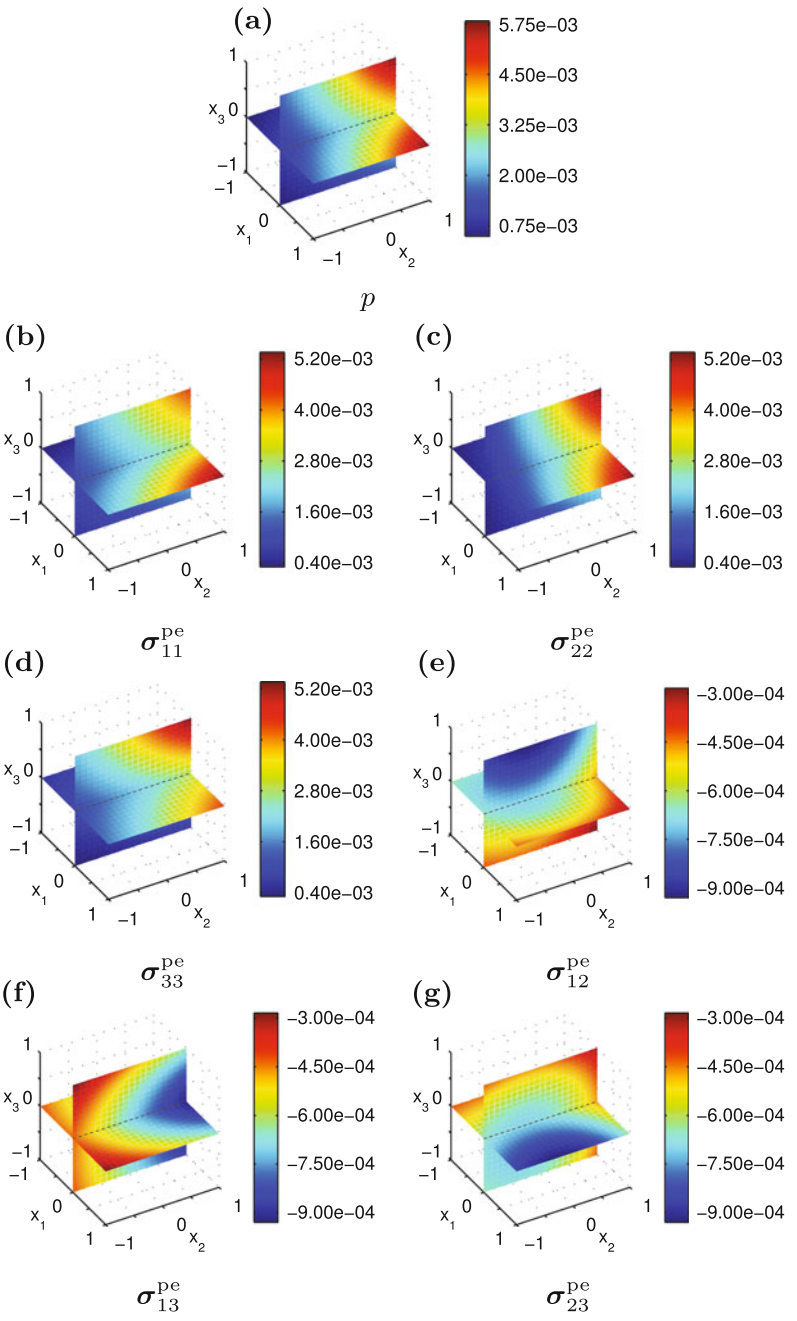

Fig. 6.62 Colored surface plots of p and $\boldsymbol{\sigma}^{\mathrm{pe}}$ for Example 6.11 evaluated on the cube $\Omega = (-1, 1)^3$ at time $t = 1$. Please note that different scales are used for different quantities as indicated by the different scales on the colorbars. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

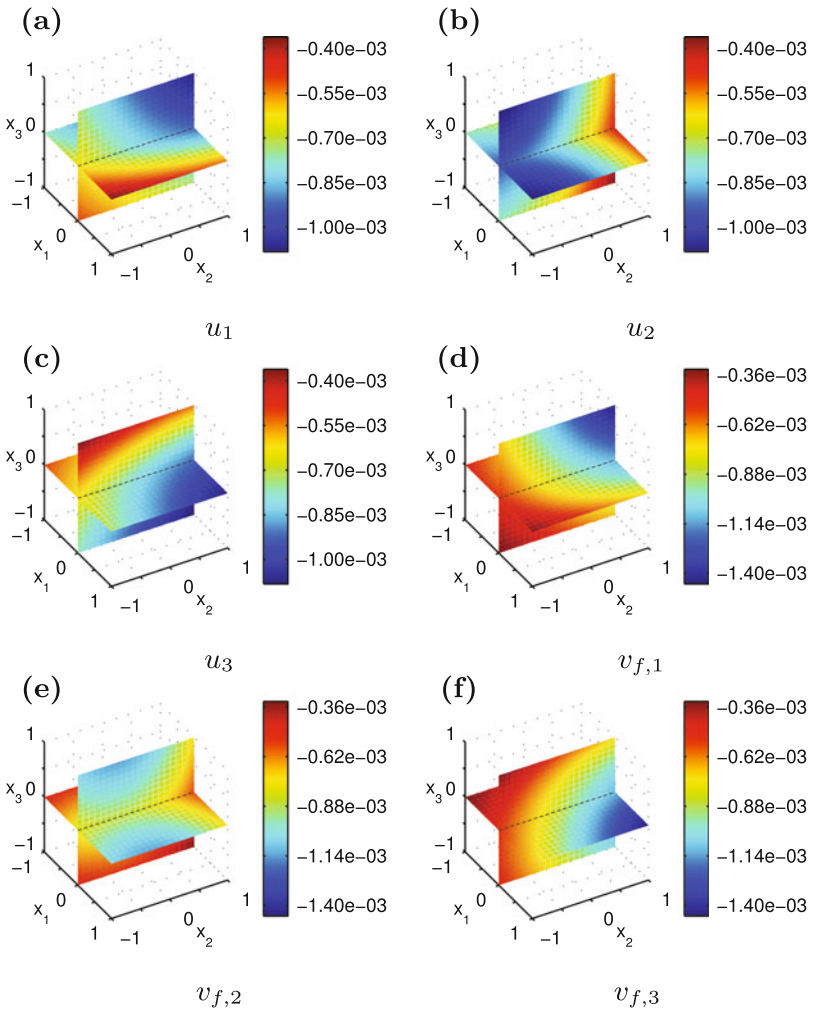

Fig. 6.63 Colored surface plots of u and v_f for Example 6.11 evaluated on the cube $\Omega = (-1, 1)^3$ at time $t = 5$. Please note that different scales are used for different quantities as indicated by the different scales on the colorbars. All quantities are dimensionless

As computations on a three-dimensional domain are more expensive in terms of computational resources, we choose $t_{\text{end}} = 1$. Three combinations of parameters are considered:

(I) $\Delta_x = \frac{1}{3}$ $(I = 6)$, $\Delta_y = \frac{1}{2}$ $(M = 4)$, $\Delta_t = 0.05$ $(J = 20)$,
(II) $\Delta_x = \frac{1}{3}$ $(I = 6)$, $\Delta_y = \frac{1}{2}$ $(M = 4)$, $\Delta_t = 0.02$ $(J = 50)$,
(III) $\Delta_x = \frac{1}{5}$ $(I = 10)$, $\Delta_y = \frac{1}{3}$ $(M = 6)$, $\Delta_t = 0.05$ $(J = 20)$.

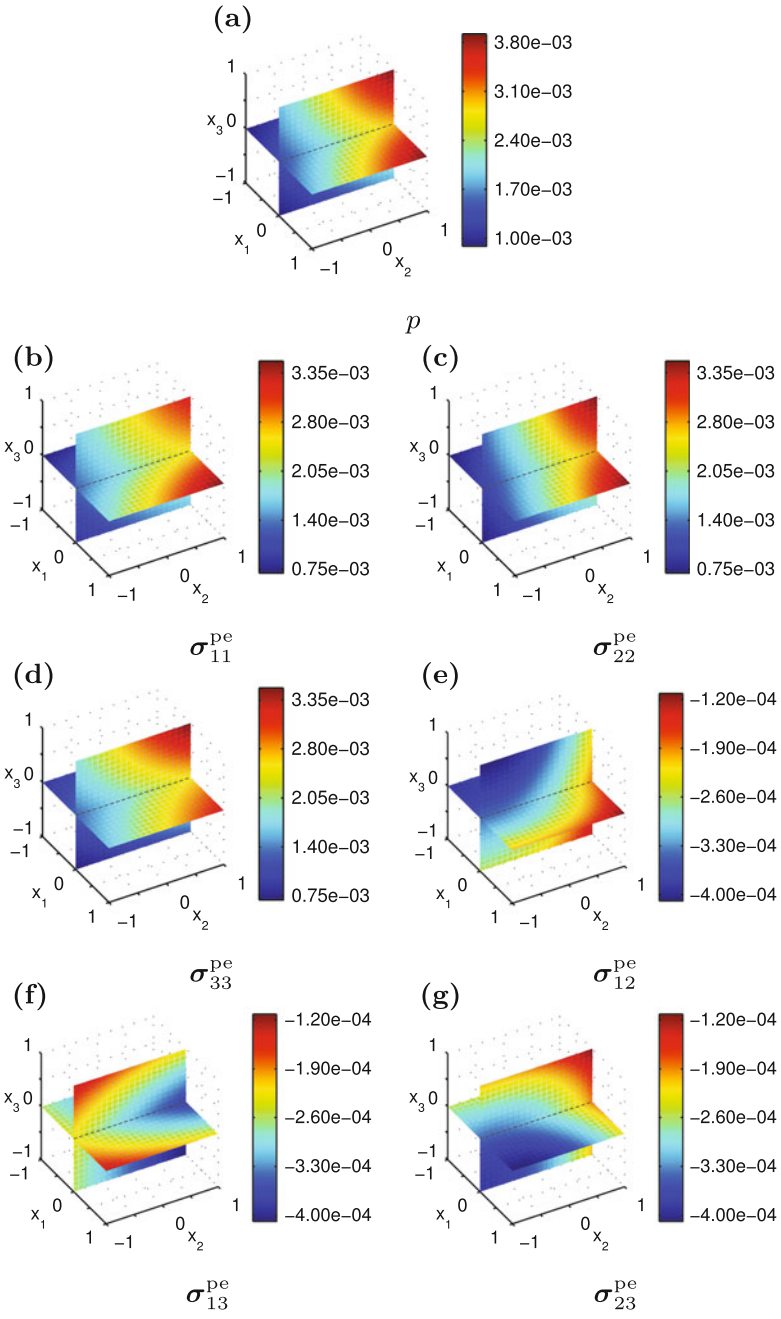

Fig. 6.64 Colored surface plots of p and $\boldsymbol{\sigma}^{\mathrm{pe}}$ for Example 6.11 evaluated on the cube $\Omega = (-1, 1)^3$ at time $t = 5$. Please note that different scales are used for different quantities as indicated by the different scales on the colorbars. In particular, the scale for the pressure p has a different sign. All quantities are dimensionless

This choice is led by the expectations that parameters yielding good results in a two-dimensional setting are good candidates for performing well in a three-dimensional setting. The other parameters are $\Delta_\tau = \Delta_t$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$. γ is varied in steps of 0.5 between 0.5 and 5. For comparison, the obtained matrices $\boldsymbol{A}$ have dimensions 21,376 × 9,600 with a condition number of $2.8 \cdot 10^{17}$, 53,848 × 24,000 with a condition number of 5.1×10^{17}, and 58,648 × 21,600 with a condition number of $5.6 \cdot 10^{17}$ for combinations (I), (II), and (III), respectively. Despite the increased expenses, we decided to use singular value decomposition to benefit from its robustness.

For the computation of the rooted mean square error with respect to the domain, $\mathrm{E}_{\mathrm{rms,x}}^{\mathrm{rel}}$, we use the grid

$$\begin{aligned} \mathfrak{Z}_{3\mathrm{D}} = \quad & \{x_1 = -1 + 0.1\,i, i \in \mathbb{N}_0, i \leq 20\} \\ & \times \{x_2 = -1 + 0.1\,j, j \in \mathbb{N}_0, j \leq 20\} \\ & \times \{x_3 = -1 + 0.1\,k, k \in \mathbb{N}_0, k \leq 20\}, \end{aligned} \tag{6.25}$$

with $I_{\mathrm{tot}} = 9{,}261$ in (6.10d). It is evaluated at the times given by $\mathfrak{Z}_t^{(1)}$ in (6.19). The rooted mean square error $\mathrm{E}_{\mathrm{rms,t}}^{\mathrm{rel}}$ with respect to time is only evaluated at the origin $x = 0$. These errors, together with the maximum relative errors as well as the points of $\mathfrak{Z}_{3\mathrm{D}}$ and times of $\mathfrak{Z}_t^{(1)}$ at which they are obtained are given in Tables 6.12–6.14.

As we can see, errors are at most about one order of magnitude larger as in the examples on the square $(-1, 1)^2$ when using the same parameter combinations. Thus, they are still very small. Varying γ between 0.5 and 5 did not yield better results than $\gamma = 1$.

Comparing the results, we see that the largest errors are in all cases achieved at the boundary of the cube, in agreement with results on the square. However, there is a difference when we consider the time at which those errors occur. As we see in Tables 6.12 and 6.13, the largest relative errors for $\Delta_x = \frac{1}{3}$ and $\Delta_y = \frac{1}{2}$ are found for small times. In contrast to this, Table 6.14 shows us that for the parameter combination with $\Delta_x = \frac{1}{5}$ and $\Delta_y = \frac{1}{3}$, the largest errors are given at the end of the time interval. This may be due to a better approximation of the initial condition in the latter case, yielding smaller errors for small times. It is interesting that decreasing Δ_t does not change where in time the largest errors are found.

The results presented in Tables 6.12 and 6.14 support our claim that parameter combinations with a good performance in a two-dimensional setting also achieve small approximation errors in a three-dimensional setting. Furthermore, we see once again that the influence of spatial parameters is stronger than the influence of temporal parameters for positive time. Nevertheless, a thorough investigation of the MFS in quasistatic poroelasticity is needed, but omitted here due to limitations on the available computational power and time.

Table 6.12 Errors in the poroelastic quantities for Example 6.11. Parameters are $\Delta_x = \frac{1}{3}$ ($I = 6$), $\Delta_y = \frac{1}{2}$ ($M = 4$), $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 1$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}(0,0,0)$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as well as the maximum relative error $\mathrm{E}^{\text{rel}}_{\text{max}}$ as defined in (6.15), (6.16), and (6.10b), and points of $\mathfrak{Z}_{3\mathrm{D}}$ as well as times of $\mathfrak{Z}_t^{(1)}$ at which they are obtained

Quantity	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}(0,0,0)$	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t	$\mathrm{E}^{\text{rel}}_{\text{max}}$	x_1	x_2	x_3	t
u_1	6.4374e–09	2.1550e–08	0.2	1.4745e–07	−0.6	0.4	1.0	0.2
u_2	6.4998e–09	2.2012e–08	0.2	1.5373e–07	0.5	−0.6	−1.0	0.2
u_3	5.0066e–08	8.6427e–08	0.2	6.0781e–07	−0.3	−0.7	−0.6	0.2
p	5.2452e–08	2.6162e–08	0.2	7.6946e–07	−1.0	−1.0	−1.0	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{11}$	2.0276e–07	4.0505e–07	0.2	5.4669e–06	−1.0	1.0	−0.2	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{22}$	2.0726e–07	4.0755e–07	0.2	5.7177e–06	1.0	−1.0	−0.3	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{33}$	2.1344e–07	2.8164e–07	0.4	1.6528e–06	−1.0	−1.0	−0.1	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{12}$	7.4230e–08	1.0025e–06	0.2	7.5029e–06	1.0	1.0	−0.8	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{13}$	1.3873e–08	4.4501e–07	0.4	3.3606e–06	1.0	−1.0	1.0	0.4
$\boldsymbol{\sigma}^{\text{pe}}_{23}$	7.3641e–09	4.4790e–07	0.4	3.2870e–06	−1.0	1.0	1.0	0.4
$v_{f,1}$	1.4427e–08	1.2118e–07	0.2	5.8506e–06	1.0	−1.0	−1.0	0.2
$v_{f,2}$	1.7058e–08	1.1934e–07	0.2	5.7459e–06	−1.0	1.0	−1.0	0.2
$v_{f,3}$	4.8725e–08	7.8054e–08	0.2	2.3510e–06	−1.0	−1.0	−0.9	0.2

Table 6.13 Errors in the poroelastic quantities for Example 6.11. Parameters are $\Delta_x = \frac{1}{3}$ ($I = 6$), $\Delta_y = \frac{1}{2}$ ($M = 4$), $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.02$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 1$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}(0,0,0)$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as well as the maximum relative error $\mathrm{E}^{\text{rel}}_{\text{max}}$ as defined in (6.15), (6.16), and (6.10b), and points of $\mathfrak{Z}_{3\mathrm{D}}$ as well as times of $\mathfrak{Z}_t^{(1)}$ at which they are obtained

Quantity	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}(0,0,0)$	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t	$\mathrm{E}^{\text{rel}}_{\text{max}}$	x_1	x_2	x_3	t
u_1	7.4707e–09	2.2107e–08	0.2	1.5123e–07	−0.6	0.4	−1.0	0.2
u_2	7.6915e–09	2.2376e–08	0.2	1.5120e–07	0.8	0.8	−1.0	0.2
u_3	5.2840e–08	8.2526e–08	0.2	5.5149e–07	−0.3	−0.3	1.0	0.2
p	3.2212e–08	2.5036e–08	0.2	7.4991e–07	−1.0	−1.0	−1.0	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{11}$	2.0272e–07	3.7755e–07	0.2	5.4237e–06	−1.0	1.0	−0.2	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{22}$	2.0060e–07	3.7374e–07	0.2	5.3105e–06	1.0	−1.0	−0.2	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{33}$	2.1164e–07	2.9764e–07	0.4	1.9312e–06	−1.0	−1.0	−0.1	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{12}$	3.6431e–08	9.7846e–07	0.4	7.5153e–06	1.0	1.0	−0.8	0.2
$\boldsymbol{\sigma}^{\text{pe}}_{13}$	1.1521e–08	4.5928e–07	0.4	3.1160e–06	1.0	−1.0	1.0	0.4
$\boldsymbol{\sigma}^{\text{pe}}_{23}$	9.0137e–09	4.5643e–07	0.4	3.1003e–06	−1.0	1.0	1.0	0.4
$v_{f,1}$	9.6707e–09	1.0013e–07	0.2	5.0913e–06	1.0	−1.0	−1.0	0.2
$v_{f,2}$	9.2490e–09	1.0271e–07	0.2	5.0229e–06	−1.0	1.0	−1.0	0.2
$v_{f,3}$	6.7261e–08	7.0024e–08	0.2	2.1182e–06	−1.0	−1.0	−1.0	0.2

Table 6.14 Errors in the poroelastic quantities for Example 6.11. Parameters are $\Delta_x = \frac{1}{5}$ ($I = 10$), $\Delta_y = \frac{1}{3}$ ($M = 6$), $\gamma = 1$, $\Delta_\tau = \Delta_t = 0.05$, $\tau_{-1} = \tau_1 = 0.5\Delta_t$, $t_{\text{end}} = 1$. Given are the maximum relative rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,t}}(0,0,0)$ with respect to time and the maximum rooted mean square error $\mathrm{E}^{\text{rel}}_{\text{rms,x}}$ with respect to the domain as well as the maximum relative error $\mathrm{E}^{\text{rel}}_{\text{max}}$ as defined in (6.15), (6.16), and (6.10b), and points of $\mathfrak{Z}_{3\text{D}}$ as well as times of $\mathfrak{Z}_t^{(1)}$ at which they are obtained

Quantity	$\mathrm{E}^{\text{rel}}_{\text{rms,t}}(0,0,0)$	$\mathrm{E}^{\text{rel}}_{\text{rms,x}}$	t	$\mathrm{E}^{\text{rel}}_{\text{max}}$	x_1	x_2	x_3	t
u_1	3.5885e–11	1.2454e–10	1.0	7.3608e–10	0.1	0.4	1.0	1.0
u_2	3.9809e–11	1.0179e–10	1.0	6.0759e–10	0.5	0.1	1.0	1.0
u_3	2.4586e–10	1.5823e–10	1.0	1.3965e–09	−0.4	0.1	1.0	1.0
p	5.2186e–11	1.0388e–10	1.0	1.2928e–09	−1.0	−1.0	−1.0	0.2
σ^{pe}_{11}	2.2557e–10	4.5569e–10	1.0	9.0816e–09	−1.0	−1.0	−1.0	1.0
σ^{pe}_{22}	1.1788e–10	3.6221e–10	1.0	5.5253e–09	−0.6	−1.0	−1.0	1.0
σ^{pe}_{33}	4.8630e–11	2.3872e–10	1.0	2.9708e–09	−1.0	−1.0	−0.5	1.0
σ^{pe}_{12}	8.6640e–10	1.8071e–09	1.0	1.7248e–08	1.0	1.0	0.5	0.8
σ^{pe}_{13}	2.7887e–10	1.5136e–09	1.0	1.0897e–08	1.0	0.4	−0.1	1.0
σ^{pe}_{23}	1.2750e–10	1.4105e–09	1.0	9.8990e–09	0.4	1.0	0.6	1.0
$v_{f,1}$	1.3168e–10	9.4630e–10	1.0	1.1399e–08	−1.0	−0.7	−1.0	0.2
$v_{f,2}$	1.6265e–10	9.8499e–10	1.0	1.0883e–08	−1.0	−1.0	−1.0	0.2
$v_{f,3}$	3.3164e–10	5.2389e–10	1.0	4.9507e–09	−1.0	−1.0	−0.9	0.2

6.8 Concluding Remarks

Whereas we have shown in the previous chapter that the MFS can in principle be used to calculate approximate solutions to initial boundary value problems of poroelasticity, this chapter aims at showing the actual capabilities of the resulting numerical method, but also its limits. For this purpose, we have considered a rather large number of examples. Our examples are carefully chosen, also with regard to a practitioner's point of view and interests. The analysis starts with what can probably be considered to be the least difficult setup as given in Example 6.2. Here, the solution under consideration is of the same type as the ansatz functions used to reconstruct it. Moreover, Dirichlet boundary conditions are given on the whole of the boundary, i.e., the quantities u and p that we use as primary variables are known on the whole boundary. As the results in Sect. 6.2 clearly show, the MFS can easily handle this example. The situations in Examples 6.3 and 6.4 are each slightly more difficult. The difficulty in the former is that we now have mixed boundary conditions, i.e., our method has to deal with the fact that on a part of the boundary, only quantities related to co-normal derivatives of our primary variables are known. However, the solution for which we are looking is still of the same kind as the ansatz functions. In the latter, the situation is reversed in the sense that Dirichlet boundary conditions are given on the whole boundary, but the solution that has to be reconstructed is not of the same type as the ansatz functions. Again, the results in Sect. 6.2 show that the MFS can handle such settings. Finally, both difficulties

are combined in Example 6.5 which, thus, probably gives the best insight into the performance of the MFS in quasistatic poroelasticity, which is still remarkably good.

Although the discussions on the approximation quality could be done by regarding only one of the relevant poroelastic quantities and show how different parameter choices influence the approximation of this one quantity, thus giving a rather trim depiction, we took the perhaps more cumbersome way of regarding the influence of parameter choices on the approximation error for all relevant poroelastic quantities, i.e., pressure p and all components of the displacement vector u, the stress tensor $\boldsymbol{\sigma}^{\mathrm{pe}}$ as well as the fluid velocity v_f. This is motivated by the fact that all of these quantities are of considerable importance. Different underlying questions may ask for the evaluation of different of these physically meaningful quantities. It is therefore indispensable, from the viewpoint of physics and engineering, to make sure that the MFS approach taken here is not flawed by lacking the ability to give good approximations of other quantities than the primarily used pressure p and displacement vector u. Nevertheless, we see that the parameter dependence of the approximation quality follows mostly the same patterns for all quantities under consideration. The reader interested in an overview therefore may concentrate on one subfigure of the respective figures. Probably the best choice for this is $\sigma_{11}^{\mathrm{pe}}$ (right column, second row in each figure), as it incorporates the pressure p and derivatives of the displacement vector u.

Having demonstrated the good performance of the MFS, we investigated more technical aspects. Our development of the MFS in poroelasticity was inspired by the idea of reducing demands on computational power and memory. However, the MFS results in the drawback of ill-conditioned systems of linear equations. Thus, one would assume that robust solution schemes for these systems, such as the singular value decomposition, have to be used resulting in higher computational costs. But the question whether the application of such methods is actually necessary is currently a matter of debate. The approximation errors which we achieved by using a simple least-squares method are one order of magnitude larger, but still very small. In order to actually evaluate the necessity of more sophisticated solution schemes, further investigations should consider noisy input data. On the other hand, our idea of neglecting the fi-parts of fundamental solutions to further reduce memory requirements turned out to even improve the approximation quality. Nevertheless, the memory requirements of the MFS still suffer from the fact that in our formulation, all times have to be considered at once. We failed in trying to reproduce approximation results of a time-marching scheme for the heat equation in Sect. 6.5 and could not compute any sensible approximate solution for poroelastic problems.

As mentioned above, this chapter also aims at showing the limits of the MFS. A well known problem for all numerical methods is the occurrence of steep gradients in the actual solution, yielding unsatisfying approximate solutions, often due to Gibbs phenomena. To deal with such problems, new stabilization techniques need to be incorporated in the approximation process. The MFS is no exception, as Sect. 6.6 luculently shows. However, it is unclear how such a stabilization can be applied as most techniques operate by modifying the underlying differential equation in either

its strong or weak formulation, which is automatically satisfied by any approximate solution that our approach yields, or by modifying the matrix of the system of linear equations, which is the same for examples with smooth and steep gradients and for which we have seen in examples of the former kind that even using a simple least-squares scheme produces only small errors.

For the sake of completeness, we concluded this chapter with an example on a three-dimensional domain. The very good results which we achieved in this case give an impressive demonstration of our method's capabilities with respect to dealing with applications that are closer to the challenge of modeling a real world geothermal reservoir.

Chapter 7
Conclusion and Outlook

7.1 Summary of Main Results

The aim of this thesis was the development of a numerical method to model the stress field in geothermal reservoirs. For this purpose, we first had to decide on a suitable physical model. As geothermal facilities using deep reservoirs work by extracting hot water from the reservoir through a well and re-injecting it through another well, interactions between the water pressure and stresses within the reservoir occur. Those poroelastic effects were included by using the poroelastic equations by Biot [34–38] in a form that uses the displacement vector of the solid, u, and pore pressure of the fluid, p, as primary variables. In the derivation of these equations, it is assumed that only small displacements are considered, such that the strain can be expressed by the linear infinitesimal strain. Moreover, a linear stress-strain-relation is used. For simplicity, we assumed a homogeneous and isotropic medium. The fluid flow is assumed to be governed by Darcy's law, such that the fluid velocity is proportional to the negative pressure gradient. Furthermore, Biot introduced the so-called fluid content ζ which only accounts for changes in fluid volume due to mass transport. With these assumptions, a set of poroelastic equations can be derived from balance of momentum and balance of mass. As our interest was not to model wave phenomena but consolidation processes, we used a non-dimensionalizing argument to show that terms involving second order time derivatives of u can be neglected, yielding the quasistatic equations of poroelasticity (QEP).

Assuming that the reservoir can be represented by a bounded domain Ω, we completed Biot's equations with boundary conditions and a suitable initial condition to get an initial boundary value problem in $\Omega \times [0, t_{\text{end}}]$. In this initial boundary value problem, we allowed that $u(x,t)$ and $p(x,t)$ were prescribed on one part of the boundary $\Gamma = \partial\Omega$ (Dirichlet conditions) and that normal tensions $t_n(x,t) = \boldsymbol{\sigma}^{\text{pe}}(x,t)n(x,t)$ and normal velocities $v_f(x,t) \cdot n(x,t) = -(\nabla_x p(x,t)) \cdot n(x,t)$

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4_7

(analogon to Neumann conditions) were prescribed on the complementing part of the boundary. Starting from this, we proved that a unique weak solution to the problem exists under the assumption that all parameters are positive scalar constants. For this purpose, the equations were split into one determining the displacement vector u if the pressure p is known and another one determining the pressure p. The first one reduced to the scenario encountered in linear elasticity, thus to proving coercivity of the bilinear form associated with the Cauchy-Navier equation and applying the Lax-Milgram theorem. The second one turned out to be an implicit evolution equation, therefore needing more general results by Lions to proof existence and uniqueness. As a consequence of the involved bilinear forms even being inner products, results on the regularity of p could be improved. A close inspection of the conditions on the right-hand side of the implicit evolution problem led to additional conditions on the original right-hand side. From those, an until then unknown statement on the regularity of the first time derivative of the displacement vector u could be concluded.

In the course of the above mentioned proof, we found that a suitable initial condition has to provide an initial fluid content ζ^0. As a consequence, it is not possible to choose an initial pressure and an initial displacement vector independently. Instead of providing ζ^0 directly, one may prescribe an initial pressure p^0 and boundary conditions for the displacement vector u at $t = 0$ which together determine ζ^0.

As already mentioned, we used dimensional analysis to derive a dimensionless version of the QEP. Using dimensionless equations often helps to explore which combination of material parameters generate similar behavior. Moreover, using dimensionless quantities simplifies further inspections and implementation for numerical solution approaches. Therefore, all further investigations were made with regard to the dimensionless equations.

As a first step to reduce computational costs, we decided to derive boundary integral equations to the QEP, hence reducing the dimension of the problem by one. We began by deriving the poroelastic versions of Green's first and second identities. However, as the associated differential operator is not self-adjoint, this yielded adjoint equations to the QEP. To the best of our knowledge, this is the first time that Green's identities for the QEP are derived following a strict mathematical procedure using adjoint equations instead of by incorporating a reciprocal relation.

In order to derive Green's third identity, we needed fundamental solutions to these adjoint equations, which are closely related to the fundamental solutions of the QEP. We derived those fundamental solutions based on a suggestion by Biot [37] who introduced a scalar potential for his equations. As the equations of poroelasticity consist of a scalar- and a vector-valued equation, a fundamental solution is a tensor (of rank three for two-dimensional domains, of rank four for three-dimensional domains). The close relations of Biot's equations to the heat equation, Cauchy-Navier equation, and Stokes equations are reflected in the fundamental solutions which incorporate the fundamental solutions to those other differential equations. Consequently, some properties of the fundamental solutions of quasistatic poroelasticity are inherited from those other fundamental solutions.

Incorporating the fundamental solutions of the QEP allowed us to derive Green's third identity, from which boundary integral equations were obtained. These boundary integral equations resembled well-known single- and double-layer potentials. However, we are cautious to actually call these boundary integrals the single- and double-layer of poroelasticity, as it is out of the scope of this thesis to prove limit and jump relations. To the best of our knowledge, such a proof cannot be found in the literature, yet, and it is up to further research to investigate the exact nature of limit and jump relations in quasistatic poroelasticity. Some difficulties may be expected as the fundamental solutions involve Dirac's delta distribution with respect to time. Thus, one either has to deal with these distributions or with the fact that single- and double-layer potentials incorporate summands involving integration with respect to $\Gamma \times (0, t_{\text{end}})$ and others involving only integration with respect to $\Gamma = \partial\Omega$.

Although boundary integral formulations are useful to reduce the dimension of a problem, several other difficulties arise. One of these is the fact that the emerging boundary integrals are usually singular integrals. In order to regularize them, we used an idea by Runge [239]: Instead of modeling a solution to a differential equation by an integral over the actual boundary Γ of the domain Ω, the integral is taken with respect to the boundary $\hat{\Gamma}$ of a larger domain $\hat{\Omega}$ which encloses Ω. To deal with domains that contain holes, we introduced the concept of an embracing pseudo-boundary. However, even when using this kind of regularization, an integration over the boundary is still needed. Using the Laplace equation as an example, we explained how a collocation version of the method of fundamental solutions (MFS) can be derived from Runge's regularization approach. Analogously, an MFS for the QEP was introduced and put into context by a short historical overview on the literature dealing with the MFS for a variety of differential equations. The method does not need any meshing of the domain or boundary, thus allowing a high degree of flexibility. Furthermore, as a collocation method, it does not involve any kind of integration, which is usually a procedure requiring many resources in terms of computational power, time, and memory.

In order to prove that an MFS is qualified to compute a numerical approximate solution to an initial boundary value problem in quasistatic poroelasticity, we proved that a suitable set of fundamental solutions is a dense subspace of a suitable solution space to such a problem. Assuming that an embracing pseudo-boundary $\hat{\Gamma}$ for the domain Ω is given, we chose a set of fundamental solutions such that the corresponding set of singularity points (called source points) is dense in $\hat{\Gamma} \times (0, t_{\text{end}})$. It is shown that such a set is a dense subspace of $\left(H^{-1}\left([0, t_{\text{end}}], C^l(\overline{\Omega})\right)\right)^4$ which contains all solutions to a boundary value problem to the QEP with vanishing initial conditions. For this result, Ω only has to satisfy the segment condition. $\hat{\Omega}$ should have a higher regularity, such that the boundary integrals over $\hat{\Gamma}$ are well-defined. However, as we were free in the choice of $\hat{\Omega}$, sufficient regularity of $\hat{\Gamma}$ can always be achieved.

Due to the contributions to the fundamental solutions which contain Dirac's delta distribution with respect to time, it is not possible to increase the regularity of the solution space with respect to time.

To the best of our knowledge, density results for systems containing a dependence on time, systems that cannot be decoupled into several separate equations, systems mixing equations that do and equations that do not depend on time, or even scalar-valued implicit evolution equations were not known in the literature before this thesis.

To incorporate non-vanishing initial values, a density result for the approximation of initial values for the heat equation in $L^2(\Omega)$ was shown. This is sufficient as we have seen in previous chapters that it is appropriate to prescribe initial values to the fluid content ζ which is governed by the heat equation. Motivated by the results concerning the approximation of initial values, we had a closer look at the formulation of the QEP if ζ is used as a primary variable instead of the pressure p. In this formulation, one of the equations only involves ζ without involving u, i.e., it decouples from the other one. This allowed us to derive an alternative MFS, called reduced ansatz, as it does not include all parts of the fundamental solutions incorporated in our first approach. Density results for the reduced ansatz follow from such results for the heat equation and Cauchy-Navier equation. The reduced ansatz further decreased the requirements on memory for the MFS. Moreover, it can directly be transferred to quasistatic thermoelasticity. Thus, it generalizes a result in [264], in which only static thermoelasticity is considered.

In order to examine the numerical performance of the MFS for quasistatic poroelasticity, we considered several exemplary problems on the square $\Omega = (-1, 1)^2$. Approximate solutions were computed on Ω using the reduced ansatz. Despite an extensive literature research, we could not find that the MFS has been applied to a system of equations that is non-elliptic before. Thus, applying it to a system of equations of mixed type, one of which shows an explicit dependence on time, is a novelty on more than one account.

As not many analytical solutions to the QEP are known, and most of those require the special case that some material parameters vanish, we considered different parts of the fundamental solutions tensor for which we placed the singularity well outside Ω and at negative time $\tau = -1$. The corresponding analytical solutions on $(-1, 1)^2$ were rather smooth. The considered initial boundary value problems prescribed either p and u on the whole boundary (pure Dirichlet conditions) or just on a part of the boundary, whereas normal tension and normal fluid velocities were prescribed on the complementary boundary parts (mixed conditions). We used these examples to investigate the influence of different groups of parameters. The first group under consideration, called spatial parameters, included the number of collocation points at the boundary Γ (corresponding to parameter Δ_x), number of source points on the pseudo-boundary $\hat{\Gamma}$ (corresponding to parameter Δ_y), and the distance γ between domain boundary Γ and pseudo-boundary $\hat{\Gamma}$. It appeared that certain combinations of these parameters provide much better results than others. In particular, some combinations with more collocation than source points, therefore yielding an overdetermined system of linear equations, with source points chosen close to the boundary, achieved the smallest approximation errors. This is in contrast to other results for the MFS in the context of elliptic equations, for which

underdetermined systems performed better or the source points had to be chosen at larger distances to the domain boundary.

The second set included the absolute value of the largest negative source point time, τ_{-1}, and the temporal distance between source points located at negative times, Δ_τ. Compared to variations of the spatial parameters, variations in the parameters of the second set had a much smaller influence on the numerical performance. Exceptions were choices with large Δ_τ and combinations of small Δ_τ with small τ_{-1}. This behavior could be expected from known results and is probably due to the fact that those parameter choices yield many similar columns in the matrix of the resulting system of linear equations. The considerations of these sections complete previous results on the placement of source points with negative τ and provide a better understanding of the reasons for poorer performance of specific parameter choices.

The third set of parameters consisted of the temporal distance between collocation points Δ_t and the smallest positive time at which source points were located, τ_1. As with the second set, the influence of the third set on the performance of the method was rather small. Variations of τ_1 only had a very minor influence. Decreasing Δ_t from the value $\Delta_t = 0.05$ which we also used before when varying parameters of the other two sets did not improve results much. However, increasing Δ_t too much yielded increased errors, which is probably due to the fact that the number of collocation points is much decreased for large Δ_t.

The influence of the temporal parameters Δ_t, Δ_τ, τ_{-1}, and τ_1 may be stronger if the spatial parameters Δ_x, Δ_y, and γ are chosen suboptimal. This was not tested here as it is probably always possible to optimize the spatial location of source points for given collocation points first. Moreover, the temporal parameters have no influence on the placement of source points corresponding to CN- and St-parts of the fundamental solutions. This supports the claim that positions of source points should be optimized first with regard to spatial coordinates.

Concerning the distribution of relative errors in time or in the domain, we found that the largest errors are in general found at the boundary of the domain for small times. This result may have been surprising at first glance as we only considered data corresponding to boundary and initial values, but could be explained if all solutions to the QEP tend toward the same behavior with increasing time. Such a behavior would then also be shared by the fundamental solutions, yielding better approximation properties as time progresses. Especially if the exponential damping known from parabolic differential equations also occurs in quasistatic poroelasticity, the influence of the initial conditions would rapidly decrease over time, explaining the improvements in approximation with increasing time as the errors in the approximation of the initial conditions would also be damped.

Bearing in mind that the systems of linear equations which arise in the context of the MFS usually incorporate matrices with large condition numbers, we decided to use the singular value decomposition as a numerically stable solution method. Comparing those results to the ones obtained by using a simple Gaussian least-squares method, we found that the latter also performed very well with errors being around one order of magnitude larger, i.e., still in the range $5 \cdot 10^{-6}$ to $1 \cdot 10^{-10}$.

As we derived two versions of the MFS, called full and reduced ansatz, any investigation would be incomplete without a comparison of these two approaches. We decided to solve both problems with mixed boundary conditions with the full ansatz for well performing Δ_x, Δ_t, Δ_τ, τ_1, and τ_{-1} with different combinations of the parameters Δ_y and γ. Allowing the latter two parameters to vary seemed necessary for a fair comparison as we have seen that an appropriate combination of the spatial parameters Δ_x, Δ_y, and γ is crucial to obtain good results. It turned out that the full ansatz, which already increases the computational costs by involving more ansatz functions, needs a larger number of source points to achieve a good approximation. Still, the minimal errors for this ansatz were about two orders of magnitude larger than the errors for the reduced ansatz. Consequently, the reduced ansatz is preferable.

Further reduction of the computational costs would be possible if some kind of time-marching could be used. This would allow to solve a smaller system of linear equations for each time step separately instead of solving one large system regarding all collocation and source points for all times. Following an idea for such a time-marching scheme, we tried different parameters and configurations of source and collocation points. However, none of those yielded acceptable results. It is not entirely clear why this time-marching scheme, which was successfully applied to an MFS for the heat equation, failed in our case. It is known that the scheme is very sensitive to the location of collocation and source points. Thus, further variations on the choice of collocation and source points or implementation of an optimization algorithm for an adaptive source points choice may achieve better results.

Until now, only solutions which changed slowly with respect to time and space were considered. As it is well known that problems containing steep gradients are often harder to solve, we modified our previous examples such that the analytical solutions contained steep gradients in $\Omega \times (0, t_{\text{end}})$. The errors for those examples were larger, ranging between 1 and 900 % for rooted mean square errors or even being close to 10^5 for maximum relative errors. These results had been expected, as examples with steep gradients often require sophisticated numerical schemes using regularization or additional stabilizing terms. In the case of the MFS, we can assume that an adaptive choice of source points and the utilization of regularizing techniques for the underlying system of linear equations would improve our results. The implementation of such an adaptive scheme was out of the scope of this thesis and, thus, is an interesting topic for further research.

Our considerations of numerical performance of the MFS was concluded by investigating an initial boundary value problem on the cube $(-1, 1)^3$, i.e., a three-dimensional domain. Generalizing the implementation from two to three spatial dimensions was straight forward. As an example problem, we again considered one column of the fundamental solutions tensor with its singularity outside of the cube and for negative time $\tau = -1$. As boundary conditions, we assumed Dirichlet boundary conditions on four faces of the cube and a given normal tension and normal fluid velocity on the remaining two faces. We considered three different parameter combinations. Combinations (I) and (II) differed in the temporal distance of the collocation points and, thus, the temporal difference of the source points for

positive times. Combinations (I) and (III) differed in the choice of the numbers of collocation and source points. For those choices, we let ourselves be guided by our previous results on the square $(-1, 1)^2$. In order to reduce the numerical cost, we considered a smaller time interval. The approximate numerical solutions which we computed showed errors with an order of magnitude between 10^{-6} and 10^{-11} in good agreement with the results on the square. In comparison, errors on the square were about one order of magnitude smaller. The behavior of the errors with regard to the different parameter choices was also similar to the behavior on the square. Moreover, the results could not be improved by changing the distance between the domain boundary and the pseudo-boundary. This supported the claim that the parameter choices with good performance in a two-dimensional setting also achieve small approximation errors in a three-dimensional setting. Additionally, we once again observed that the influence of the spatial parameters is larger than the influence of the temporal parameters for positive times. However, a thorough investigation of the MFS for the QEP on three-dimensional domains is needed, but was omitted here due to limited computational resources and time available during the work on this thesis.

Concluding the results of this thesis with respect to the fourth column of the Kaiserslautern Model, cf. Fig. 1.2, we can say that we succeeded in the development of a numerical method to model the stress field in a geothermal reservoir. The method is applicable to three-dimensional domains, including poroelastic effects. As a meshfree method, it is flexible with respect to the location of data points. Due to the reduction of the dimension of the problem by one and, as a collocation method, being integration-free, requirements in terms of memory and computational power are decreased. If heat transport can be treated by the heat equation, i.e., convective heat transport can be neglected, it is easy to incorporate thermoelastic effects. Thus, the MFS provides a numerical solution scheme incorporating the first two items of the fourth column. As the other items are either concerned with fracture mechanics or wave phenomena, they are naturally out of the scope of the quasistatic equations of poroelasticity.

7.2 Future Research Perspectives

Throughout this thesis, we already mentioned some perspectives on further research.

In order to derive the MFS, we derived integral boundary equations similar to single- and double-layer potentials known from potential theory [91]. These layer potentials have characteristic limit and jump relations. Similar relations can be derived for the heat equation [57]. However, regarding the relations between the fundamental solutions to the QEP and fundamental solutions to the heat equation, Cauchy-Navier equation and Stokes equations, limit and jump relations in quasistatic poroelasticity may differ from the results known in potential theory.

Proving limit and jump relations in the context of quasistatic poroelasticity may also help with deriving convergence results for the MFS.

One of our goals in developing the MFS for the QEP was to reduce computational costs. It may be up to debate whether this was successful. Although we could reduce the dimension of the problem by one and avoid performing numerical integration, this comes at the price of having to consider all data for all points in time at once instead of incorporating a time-marching scheme that only needs to consider data for two consecutive points in time at once. Our attempt on applying a time-marching scheme already used for the MFS for the heat equation failed. However, further investigations are needed to determine whether this scheme is either not applicable to the QEP in principle or just needs a better suited choice of collocation and source points.

There is another possibility to apply a time-marching scheme together with the MFS. A finite difference scheme can be applied first to the QEP, yielding a new set of equations which are not directly dependent on time. The MFS can then be applied to these new equations, combining both approaches to get a solution to an initial boundary value problem. This method was already successfully applied to the heat equation [135, 281], resulting in an MFS for some kind of Helmholtz equation, as well as diffusion and convection-diffusion equations [219]. The problem of applying time-marching schemes is that they make it necessary to deal with non-vanishing right-hand sides.

By construction, the MFS – or Trefftz methods in general – are only suitable for homogeneous problems, i.e., such problems for which the right-hand side vanishes. In order to extend their applicability to inhomogeneous problems, they have to be combined with other methods. An example of such an extension is the usage of the dual reciprocity method [115, 208]. Assume that we have two systems of radial basis functions $\{u_n^{\mathrm{rad}}\}_{n=0}^N$ and $\{f_n^{\mathrm{rad}}\}_{n=0}^N$ such that

$$a(u_n^{\mathrm{rad}}, v) = f_n^{\mathrm{rad}}(v) \quad \forall v \in V \quad \forall n \in \mathbb{N}_0, 0 \leq n \leq N \,. \tag{7.1}$$

Then by finding an approximation $f^{\mathrm{rad}} = \sum_{n=1}^N b_n f_n^{\mathrm{rad}}$, the function $u^{\mathrm{rad}} = \sum_{n=1}^N b_n u_n^{\mathrm{rad}}$ is an approximate solution to (7.1), as long as no boundary condition is imposed. If a Dirichlet boundary condition with boundary values given by g is imposed, a solution can be found by combining u^{rad} and the solution of a Trefftz method approach to the differential equation under consideration with the modified boundary data $g^{\mathrm{rad}} = g - u^{\mathrm{rad}}|_\Gamma$.

In order to use radial basis functions in quasistatic poroelasticity, they have to be combined into tensors of rank three for a two-dimensional setting and rank four for a three-dimensional setting. The reader is referred to, e.g., [180, 181] for results on tensor-valued radial basis functions and [42, 77] for applications of radial basis functions for differential equations in general. Due to the coupling of pressure and displacement vector, it is not clear how a suitable set of radial basis functions may be chosen. Moreover, those functions would also have to be modified to account for time-dependence.

Alternatively, an ansatz with time-independent radial basis functions may be used to reduce the QEP to a set of ordinary differential equations with respect to time by

calculating the spatial derivative of the ansatz analytically. The resulting ordinary differential equation can then be solved with well-known methods [268].

Independent of how the problem of non-vanishing right-hand sides may be conquered, when applying a time-marching scheme it may become necessary to consider how a CFL condition looks like for the QEP (cf. Remark 6.7).

Another possibility to eliminate the time-dependence is the application of a Laplace transformation with respect to time as in [288]. For this, Laplace transforms of boundary and initial conditions would have to be computed. The problem then needs to be solved in the Laplace domain for a certain set of transformation parameters in order to provide the necessary data for a numerical backwards transformation. Although the Laplace transform of the fundamental solutions is known, the backward transformation may turn out difficult due to the involvement of Dirac's delta distribution with respect to time and its Laplace transform. Moreover, computing Laplace transforms of initial and boundary data may turn out to be difficult, especially if they are only given on a discrete set of points. Especially non-vanishing initial conditions may yield difficulties, as they would become involved in the Laplace transformation of time derivatives.

The incorporation of adaptive source point choices is certainly an important aspect in the future development of the MFS in poroelasticity. Such adaptive schemes should include the possibility to locate more collocation and source points near corners, cusps, or in general at and near parts of the boundary with larger curvature. Furthermore, the number of collocation and source points in proximity to the expected location of steep gradients should be increased. Also, source points should probably be positioned closer to the boundary in such a situation. The reader is referred to [24] for an example of an adaptive method on how to choose the locations of source points for the Helmholtz equation.

Incorporating regularization techniques may also improve the performance of the MFS and is often done due to the poor condition of the underlying system of linear equations. Although we achieved good results without regularization here, it is certainly necessary when dealing with real-world data, especially as most real-world data is superimposed with noise [29, 30, 212]. Noisy data is much more sensitive to ill-conditioning. Thus, applying the MFS to noisy data without regularization would probably not yield any acceptable results. Besides well known techniques like truncated singular value decomposition or Thikhonov regularization [212], regularization may also be achieved by using approximate fundamental solutions [235].

Long term goals would be the incorporation of thermoelastic effects (see [207] for experimental results on the influence of cold water injections on the pressure in geothermal wells), viscoelastic effects or even effects of plastic deformation. The reader is referred to [198] for a finite element approach in thermoviscoplasticity involving operator splitting. Furthermore, in order to derive a rather complete model of a geothermal reservoir, fluid and heat flow models have to be coupled with a stress field model [176, 286]. Interim stages towards this goal, or even goals in their own respect, are the incorporation of heterogeneous media, fractured media, mechanics of fracture incurrence and fracture growth, prediction and prevention of seismic

events, and incorporation of parameters that depend on the quantities which are to be computed, especially the pressure. In order to regard such parameter dependencies and still apply an MFS, some iterative procedure has to be developed as the resulting equations are non-linear. Therefore, it is no longer possible to derive fundamental solutions in closed form.

Appendix A
Fundamental Stresses and Fluid Velocities

This appendix lists the components of the stress tensors $\boldsymbol{\sigma}^{\mathrm{pe},\cdot}$ and fluid velocities $v_f^{\cdot}$ to the dimensionless fundamental solutions.

For three–dimensional domains, these are given by

$$\begin{aligned}
&\boldsymbol{\sigma}_{kl}^{\mathrm{pe,Si}}(x,t)\\
&=\frac{C_1}{2\pi\,\|x\|^3}\Bigg[\left(\delta_{kl}-\frac{3x_kx_l}{\|x\|^2}\right)\left(\operatorname{erf}\left(\frac{\|x\|}{\sqrt{4C_2t}}\right)-\frac{2}{\sqrt{\pi}}\frac{\|x\|}{\sqrt{4C_2t}}\exp\left(-\frac{\|x\|^2}{4C_2t}\right)\right)\\
&\qquad+\frac{4}{\sqrt{\pi}}\left(\delta_{kl}+\frac{x_kx_l}{\|x\|^2}\right)\frac{\|x\|^3}{\sqrt{4C_2t^3}}\exp\left(-\frac{\|x\|^2}{4C_2t}\right)\Bigg],
\end{aligned}\tag{A.1a}$$

$$\begin{aligned}
&\boldsymbol{\sigma}_{jkl}^{\mathrm{pe,fi}}(x,t)\\
&=-\frac{C_1^2}{2\pi\,\|x\|^5}\Bigg[\left(x_j\delta_{lk}+x_k\delta_{jl}+x_l\delta_{jk}-\frac{5x_jx_kx_l}{\|x\|^2}\right)\\
&\qquad\times\left(3\operatorname{erf}\left(\frac{\|x\|}{\sqrt{4C_2t}}\right)-\frac{2}{\sqrt{\pi}}\left(3+\frac{\|x\|^2}{2C_2t}\right)\frac{\|x\|}{\sqrt{4C_2t}}\exp\left(-\frac{\|x\|^2}{4C_2t}\right)\right)\\
&\qquad+\frac{8}{\sqrt{\pi}}\left(x_j\delta_{kl}-\frac{x_jx_kx_l}{\|x\|^2}\right)\frac{\|x\|^5}{\sqrt{4C_2t^5}}\exp\left(-\frac{\|x\|^2}{4C_2t}\right)\Bigg],
\end{aligned}\tag{A.1b}$$

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4

$$
\begin{aligned}
&\boldsymbol{\sigma}_{jkl}^{\mathrm{pe,CN}}(x) \\
&= \frac{-(2c_0\lambda + \alpha^2)}{c_0(\lambda + 2\mu) + \alpha^2} \frac{x_j}{4\pi \left\|x\right\|^3} \delta_{kl} \\
&\quad - \frac{C_3}{4\pi \left\|x\right\|^3} \left[(1 - C_4)(x_k\delta_{jl} + x_l\delta_{jk}) - 2C_4 x_j \delta_{kl} + 4C_4 \frac{x_j x_k x_l}{\left\|x\right\|^2} \right] ,
\end{aligned} \tag{A.1c}
$$

$$
v_{f,k}^{\mathrm{Si}}(x,t) = \frac{2C_2}{\sqrt{\pi}^3} \frac{x_k}{\sqrt{4C_2 t}^5} \exp\left(-\frac{\left\|x\right\|^2}{4C_2 t} \right) , \tag{A.2a}
$$

$$
v_{f,jk}^{\mathrm{fi}}(x,t) = \frac{2C_1 C_2}{\sqrt{\pi}^3} \left(\delta_{jk} - \frac{x_j x_k}{2C_2 t} \right) \frac{1}{\sqrt{4C_2 t}^5} \exp\left(-\frac{\left\|x\right\|^2}{4C_2 t} \right) , \tag{A.2b}
$$

$$
v_{f,jk}^{\mathrm{CN}}(x) = \frac{C_1}{4\pi \left\|x\right\|^3} \left(-\delta_{jk} + \frac{3x_j x_k}{\left\|x\right\|^2} \right) . \tag{A.2c}
$$

For two–dimensional domains, we have

$$
\begin{aligned}
&\boldsymbol{\sigma}_{kl}^{\mathrm{pe,Si}}(x,t) \\
&= \frac{C_1}{\pi \left\|x\right\|^2} \left[\left(\delta_{kl} - \frac{2x_k x_l}{\left\|x\right\|^2} \right) \left(1 - \exp\left(-\frac{\left\|x\right\|^2}{4C_2 t} \right) \right) \right. \\
&\qquad \left. + \left(-\delta_{kl} + \frac{x_k x_l}{\left\|x\right\|^2} \right) \frac{2\left\|x\right\|^2}{4C_2 t} \exp\left(-\frac{\left\|x\right\|^2}{4C_2 t} \right) \right] ,
\end{aligned} \tag{A.3a}
$$

$$
\begin{aligned}
&\boldsymbol{\sigma}_{jkl}^{\mathrm{pe,fi}}(x,t) \\
&= \frac{2C_1^2}{\pi \left\|x\right\|^4} \left[\left(x_j\delta_{kl} + x_k\delta_{jl} + x_l\delta_{jk} - \frac{4x_j x_k x_l}{\left\|x\right\|^2} \right) \right. \\
&\qquad \times \left(\left(1 + \frac{\left\|x\right\|^2}{4C_2 t} \right) \exp\left(-\frac{\left\|x\right\|^2}{4C_2 t} \right) - 1 \right) \\
&\qquad \left. + \left(x_j\delta_{kl} - \frac{x_j x_k x_l}{\left\|x\right\|^2} \right) \frac{2\left\|x\right\|^4}{(4C_2 t)^2} \exp\left(-\frac{\left\|x\right\|^2}{4C_2 t} \right) \right] ,
\end{aligned} \tag{A.3b}
$$

$$
\begin{aligned}
&\boldsymbol{\sigma}_{jkl}^{\mathrm{pe,CN}}(x) \\
&= \frac{-(2c_0\lambda + \alpha^2)}{c_0(\lambda + 2\mu) + \alpha^2} \frac{x_j}{2\pi \left\|x\right\|^2} \delta_{kl} \\
&\quad - \frac{C_3}{2\pi \left\|x\right\|^2} \left[(1 - C_4)(x_k\delta_{jl} + x_l\delta_{jk}) - 2C_4 x_j\delta_{kl} + 4C_4 \frac{x_j x_k x_l}{\left\|x\right\|^2} \right] ,
\end{aligned} \tag{A.3c}
$$

$$
v_{f,k}^{\mathrm{Si}}(x,t) = \frac{2C_2}{\pi} \frac{x_k}{(4C_2t)^2} \exp\left(-\frac{\left\|x\right\|^2}{4C_2t} \right) , \tag{A.4a}
$$

$$
v_{f,jk}^{\mathrm{fi}}(x,t) = \frac{2C_1C_2}{\pi} \left(\delta_{jk} - \frac{x_j x_k}{2C_2t} \right) \frac{1}{(4C_2t)^2} \exp\left(-\frac{\left\|x\right\|^2}{4C_2t} \right) , \tag{A.4b}
$$

$$
v_{f,jk}^{\mathrm{CN}}(x) = \frac{C_1}{2\pi \left\|x\right\|^2} \left(-\delta_{jk} + \frac{2x_j x_k}{\left\|x\right\|^2} \right) . \tag{A.4c}
$$

References

1. Abeyratne, M.K., Freeden, W., Mayer, C.: Multiscale deformation analysis by Cauchy-Navier wavelets. J. Appl. Math. **12**, 605–645 (2003)
2. Abousleiman, Y., Cheng, A.H.D., Cui, L., Detournay, E., Roegiers, J.C.: Mandel's problem revisited. Géotechnique **46**, 187–195 (1996)
3. Adams, R.: Sobolev Spaces. Academic, New York (1975)
4. Aleksidze, M.A.: On approximation solutions of a certain mixed boundary value problem in the theory of harmonic functions. Differ. Equ. **2**, 515–518 (1966)
5. Allis, R., Bromley, C., Currie, S.: Update on subsidence at the Wairakei-Tauhara geothermal system, New Zealand. Geothermics **38**, 169–180 (2009)
6. Allis, R.G., Zhan, X.: Predicting subsidence at Wairakei and Ohaaki geothermal fields, New Zealand. Geothermics **29**, 479–497 (2000)
7. Altay, C.A., Dökmeci, M.C.: A uniqueness theorem in Biot's poroelasticity theory. J. Appl. Math. Phys. **49**, 838–846 (1998)
8. Alves, C.J.S.: On the choice of source points in the method of fundamental solutions. Eng. Anal. Bound. Elem. **33**, 1348–1361 (2009)
9. Alves, C.J.S., Antunes, P.R.S.: The method of fundamental solutions applied to the calculation of eigensolutions for 2D plates. Int. J. Numer. Method. Eng. **77**, 177–194 (2009)
10. Alves, C.J.S., Silvestre, A.L.: Density results using Stokeslets and a method of fundamental solutions for the Stokes equations. Eng. Anal. Bound. Elem. **28**, 1245–1252 (2004)
11. Ansorge, R., Sonar, T.: Mathematical Models of Fluid Dynamics. Wiley, Weinheim (2009)
12. António, J., Tadeu, A., Godinho, L.: A Three-dimensional acoustics model using the method of fundamental solutions. Eng. Anal. Bound. Elem. **32**, 525–531 (2008)
13. Augustin, M.: Mathematical aspects of stress field simulations in deep geothermal reservoirs. Schr. Funkt.Anal. Geomath. **50**, 1–26 (2011)
14. Augustin, M.: On the role of poroelasticity for modeling of stress fields in geothermal reservoirs. Int. J. Geomath. **3**, 67–93 (2012)
15. Augustin, M., Bauer, M., Blick, C., Eberle, S., Freeden, W., Gerhards, C., Ilyasov, M., Kahnt, R., Klug, M., Michel, I., Möhringer, S., Neu, T., Nutz, H., I., Punzi, A.: Modeling deep geothermal reservoirs: recent advances and future perspectives. In: W. Freeden, Z. Nashed, T. Sonar (eds.) Handbook of Geomathematics, 2nd edn. Springer, New York (2015). Accepted for publication
16. Augustin, M., Caiazzo, A., Fiebach, A., Fuhrmann, J., John, V., Linske, A., Umla, R.: An assesment of discretizations for convection-dominated convection-diffusion equations. Comput. Method. Appl. Mech. Eng. **200**, 3395–3409 (2011)

M.A. Augustin, *A Method of Fundamental Solutions in Poroelasticity to Model the Stress Field in Geothermal Reservoirs*, Lecture Notes in Geosystems Mathematics and Computing, DOI 10.1007/978-3-319-17079-4

17. Augustin, M., Freeden, W., Gerhards, C., Möhringer, S., Ostermann, I.: Mathematische Methoden in der Geothermie. Math. Semesterber. **59**, 1–28 (2012)
18. Auriault, J.L.: Contribution à l'étude de la Consolidation des Sols. Ph.D. thesis, L'Université Scientifique et Médicale de Grenoble (1973)
19. Auriault, J.L., Sanchez-Palencia, E.: Étude du Comportement Macroscopique d'un Milieu Poreux Saturé déformable. J. Mec. **16**, 575–603 (1977)
20. Badmus, T., Cheng, A.H.D., Grilli, S.: A Laplace-transform-based three-dimensional BEM for poroelasticity. Int. J. Numer. Method. Eng. **36**, 67–85 (1993)
21. Baisch, S., Carbon, D., Dannwolf, U., Delacou, B., Delvaux, M., Dunand, F., Jung, R., Koller, M., Martin, C., Sartori, M., Secanell, R., Vorös, R.: Deep heat mining Basel – seismic risk analysis. SERIANEX. Tech. rep., study prepared for the Departement für Wirtschaft, Soziales und Umwelt des Kantons Basel-Stadt, Amt für Umwelt und Energie (2009)
22. Balakrishnan, K., Ramachandran, P.A.: The method of fundamental solutions for linear diffusion-reaction equations. Math. Comput. Model. **31**, 221–237 (2000)
23. Barbier, E.: Geothermal energy technology and current status: an overview. Renew. Sustain. Energy Rev. **6**, 3–65 (2002)
24. Barnett, A.H., Betcke, T.: Stability and convergence of the method of fundamental solutions for Helmholtz problems on analytic domains. J. Comput. Phys. **227**, 7003–7026 (2008)
25. Barucq, H., Madaune-Tort, M., Saint-Macary, P.: Theoretical aspects of wave propagation for Biot's consolidation problem. Monogr. Semin. Mat. García de Galdeano **31**, 449–458 (2004)
26. Barucq, H., Madaune-Tort, M., Saint-Macary, P.: On nonlinear Biot's consolidation models. Nonlinear Anal. **63**, e985–e995 (2005)
27. Barucq, H., Madaune-Tort, M., Saint-Macary, P.: Some existence-uniqueness results for a class of one-dimensional nonlinear Biot models. Nonlinear Anal. **61**, 591–612 (2005)
28. Barucq, H., Madaune-Tort, M., Saint-Macary, P.: Asymptotic Biot models in porous media. Adv. Differ. Equ. **11**, 61–90 (2006)
29. Bauer, F., Gutting, M., Lukas, M.A.: Evaluation of parameter choice methods for regularization of ill-posed problems in geomathematics. In: W. Freeden, Z. Nashed, T. Sonar (eds.) Handbook of Geomathematics, 2nd edn. Springer, New York (2015). Accepted for publication
30. Bauer, F., Lukas, M.A.: Comparingparameter [sic] choice methods for regularization of ill-posed problems. Math. Comput. Simul. **81**, 1795–1841 (2011)
31. Bear, J.: Dynamics of Fluids in Porous Media. Dover, Mineola (1988)
32. Berryman, J.G.: Comparison of upscaling methods in poroelasticity and its generalizations. J. Eng. Mech. **131**, 928–936 (2005)
33. Bertani, R.: Geothermal power generation in the world 2005–2010 update report. Geothermics **41**, 1–29 (2012)
34. Biot, M.A.: Le Problème de la Consolidation des Matières Argileuses sous une Charge. Ann. Soc. Sci. Brux. **B55**, 110–113 (1935)
35. Biot, M.A.: General theory of three-dimensional consolidation. J. Appl. Phys. **12**, 151–164 (1941)
36. Biot, M.A.: Theory of elasticity and consolidation for a porous anisotropic solid. J. Appl. Phys. **26**, 182–185 (1955)
37. Biot, M.A.: General solutions of the equations of elasticity and consolidation for a porous material. J. Appl. Mech. **78**, 91–96 (1956)
38. Biot, M.A.: Mechanics of deformation and acoustic propagation in porous media. J. Appl. Phys. **33**, 1482–1498 (1962)
39. de Boer, R.: Theoretical poroelasticity – a new approach. Chaos Solitons Fractals **25**, 861–878 (2005)
40. Bogomolny, A.: Fundamental solutions method for elliptic boundary value problems. SIAM J. Numer. Anal. **22**, 644–669 (1985)
41. Browder, F.E.: Approximation by solutions of partial differential equations. Am. J. Math. **84**, 134–160 (1962)

42. Buhmann, M.D.: Radial basis functions: theory and implementations. In: Cambridge Monographs on Applied and Computational Mathematics, vol. 12. Cambridge University Press, Cambridge (2003)
43. Chantasiriwan, S.: Methods of fundamental solutions for time-dependent heat conduction problems. Int. J. Numer. Method. Eng. **66**, 147–165 (2006)
44. Chantasiriwan, S., Johansson, B.T., Lesnic, D.: The method of fundamental solutions for free surface Stefan problems. Eng. Anal. Bound. Elem. **33**, 529–538 (2009)
45. Chen, C.S., Karageorghis, A., Smyrlis, Y.S. (eds.): The method of fundamental solutions – a meshless method. Dynamic Publishers, Atlanta (2008)
46. Chen, C.S., Rashed, Y.F., Golberg, M.A.: A Mesh-Free Method for Linear Diffusion Equations. Numer. Heat Transf. B **33**, 469–486 (1998)
47. Chen, C.W., Young, D.L., Tsai, C.C., Murugesan, K.: The method of fundamental solutions for inverse 2D Stokes problems. Comput. Mech. **37**, 2–14 (2005)
48. Chen, J.: Time domain fundamental solution to Biot's complete equations of dynamic poroelasticity. Part I: two-dimensional solution. Int. J. Solid. Struct. **31**, 1447–1490 (1994)
49. Chen, J.: Time domain fundamental solution to Biot's complete equations of dynamic poroelasticity. Part II: three-dimensional solution. Int. J. Solid. Struct. **31**, 169–202 (1994)
50. Chen, J.T., Wu, C.S., Lee, Y.T., Chen, K.H.: On the equivalence of the Trefftz method and method of fundamental solutions for Laplace and biharmonic equations. Comput. Math. Appl. **53**, 851–879 (2007)
51. Cheng, A.H.D., Detournay, E.: A direct boundary element method for plane strain poroelasticity. Int. J. Numer. Anal. Method. Geomech. **12**, 551–572 (1988)
52. Cheng, A.H.D., Detournay, E.: On singular integral equations and fundamental solutions of poroelasticity. Int. J. Solid. Struct. **35**, 4521–4555 (1998)
53. Cheng, A.H.D., Liggett, J.A.: Boundary integral equation method for linear porous-elasticity with applications to soil consolidation. Int. J. Numer. Method. Eng. **20**, 255–278 (1984)
54. Cheng, A.H.D., Predeleanu, M.: Transient boundary element formulation for linear poroelasticity. Appl. Math. Model. **11**, 285–290 (1987)
55. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
56. Ciarlet, P.G.: Mathematical Elasticity: Volume I: Three-Dimensional Elasticity. Studies in Mathematics and Its Applications. North-Holland, Amsterdam (1994)
57. Costabel, M.: Boundary integral operators for the heat equation. Integral Equ. Operat. Theory **13**, 498–552 (1990)
58. Costabel, M.: Time-dependent problems with the boundary integral equation method. In: E. Stein, R. de Borst, T.J.R. Hughes (eds.) Encyclopedia of Computational Mechanics, chap. 25. Wiley, Chichester (2004)
59. Courant, R., Friedrichs, K., Lewy, H.: Über die partiellen Differenzengleichungen der mathematischen Physik. Math. Ann. **100**, 32–74 (1928)
60. Courant, R., Friedrichs, K., Lewy, H.: On partial difference equations of mathematical physics. IBM J. Res. Dev. **11**, 215–234 (1967). Translated from German by Phyllis Fox
61. Cui, L., Abousleiman, Y.: Time-dependent poromechanical response of saturated cylinders. J. Eng. Mech. **127**, 391–398 (2001)
62. Dafermos, C.M.: On the existence and the asymptotic stability of solutions to the equations of linear thermoelasticity. Arch. Ration. Mech. Anal. **29**, 1521–1536 (1968)
63. Darcy, H.P.G.: Les Fontaines Publiques de la Ville de Dijon. Victor Dalmont, Paris (1856)
64. Dargush, G.F., Banerjee, P.K.: A time domain boundary element method for poroelasticity. Int. J. Numer. Method. Eng. **28**, 2423–2449 (1989)
65. Detournay, E., Cheng, A.H.D.: Fundamentals of poroelasticity. In: C. Fairhurst (ed.) Comprehensive Rock Engineering: Principles, Practice and Projects. Analysis and Design Method, vol. II, chap. 5, pp. 113–171. Pergamon Press, Oxford (1993)
66. Dickson, M.H., Fanelli, M.: Geothermal Energy. Earthscan, New York (2003)
67. Eberhart-Phillips, D., Oppenheimer, D.H.: Induced seismicity in The Geysers geothermal area, California. J. Geophys. Res. **89**, 1191–1207 (1984)

68. Edwards, H.M.: Advanced calculus: a differential forms approach. Birkhäuser, Boston (1994)
69. Ehrenpreis, L.: On the theory of kernels of Schwartz. Proc. Am. Math. Soc. **7**, 713–718 (1955)
70. Engel, K.J., Nagel, R.: A Short Course in Operator Semigroups. Universitext. Springer, New York (2006)
71. Evans, K.F., Cornet, F.H., Hashida, T., Hayashi, K., Ito, T., Matsuki, K., Wallroth, T.: Stress and rock mechanics issues of relevance to HDR/HWR engineered geothermal systems: review of developments during the past 15 years. Geothermics **28**, 455–474 (1999)
72. Evans, L.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19. American Mathematical Society, Providence (1998)
73. Expertengruppe "Seismisches Risiko bei hydrothermaler Geothermie": Das seismische Ereignis bei Landau vom 15. August 2009, Abschlussbericht. Tech. rep., on behalf of the Ministerium für Umwelt, Landwirtschaft, Ernährung, Weinbau und Forsten des Landes Rheinland-Pfalz (2010)
74. Fairweather, G., Karageorghis, A.: The method of fundamental solutions for elliptic boundary value problems. Adv. Comput. Math. **9**, 69–95 (1998)
75. Fairweather, G., Karageorghis, A., Martin, P.A.: The method of fundamental solutions for scattering and radiation problems. Eng. Anal. Bound. Elem. **27**, 759–769 (2003)
76. Fam, G.S.A., Rashed, Y.F.: The method of fundamental solutions applied to 3d structures with body forces using particular solutions. Comput. Mech. **36**, 245–254 (2005)
77. Fasshauer, G.E.: Solving partial differential equations by collocation with radial basis functions. In: A. Le Méhauté, C. Rabut, L.L. Schumaker (eds.) Surface Fitting and Multiresolution Methods, pp. 131–138. Vanderbilt University Press, Nashville (1997)
78. Fialko, Y., Simons, M.: Deformation and seismicity in the Coso geothermal area, Inyo County, California: observations and modeling using satellite radar interferometry. J. Geophys. Res. **105**, 21,781–21,793 (2000)
79. Freeden, W.: On the approximation of external gravitational potential with closed systems of (trial) functions. Bull. Géod. **54**, 1–20 (1980)
80. Freeden, W.: Least squares approximation by linear combination of (multi-)poles. Report 344, The Ohio State University, Departement of Geodetic Science and Surveying, Columbus (1983)
81. Freeden, W.: Multiscale Modelling of Spaceborne Geodata. Teubner, Stuttgart (1999)
82. Freeden, W.: Metaharmonic Lattice Point Theory. CRC Press/Taylor & Francis, Boca Raton (2011)
83. Freeden, W., Blick, C.: Signal decorrelation by means of multiscale methods. World Min. Surf. Undergr. **65**, 304–317 (2013)
84. Freeden, W., Gerhards, C.: Poloidal and toroidal field modeling in terms of locally supported vector wavelets. Math. Geosci. **42**, 817–838 (2010)
85. Freeden, W., Gerhards, C.: Geomathematically Oriented Potential Theory. Chapman & Hall/CRC Press, Boca Raton (2013)
86. Freeden, W., Gervens, T., Schreiner, M.: Constructive Approximation on the Sphere (With Applications to Geomathematics). Oxford Science Publications, Oxford (1998)
87. Freeden, W., Groten, E., Schreiner, M., Söhne, W., Tücks, M.: Deformation analysis using Navier spline interpolation (with an application to the Lake Blåsjø area). Allg. Vermess.-Nachr. **3**, 120–146 (1996)
88. Freeden, W., Gutting, M.: Special Functions of Mathematical (Geo-)Physics. Birkhäuser, Basel (2013)
89. Freeden, W., Kersten, H.: A constructive approximation theorem for the oblique derivative problem in potential theory. Math. Method. Appl. Sci. **3**, 104–114 (1981)
90. Freeden, W., Mayer, C., Schreiner, M.: Tree algorithms in wavelet approximations by Helmholtz potential operators. Numer. Funct. Anal. Optim. **24**, 747–782 (2003)
91. Freeden, W., Michel, V.: Multiscale Potential Theory with Applications to Geoscience. Applied and Numerical Harmonic Analysis. Birkhäuser, Boston (2004)
92. Freeden, W., Michel, V.: Wavelet deformation analysis for spherical bodies. Int. J. Wavelets, Multiresolut. Inf. Process. **3**, 523–558 (2005)

93. Freeden, W., Nutz, H.: Satellite gravity gradiometry as tensorial inverse problem. Int. J. Geomath. **2**, 123–146 (2011)
94. Freeden, W., Nutz, H.: Mathematische Methoden. In: M. Bauer, W. Freeden, H. Jacobi, T. Neu (eds.) Handbuch Tiefe Geothermie. Springer, Berlin (2014)
95. Freeden, W., Ostermann, I., Augustin, M.: Mathematik als Schlüsseltechnologie in der Geothermie. Geotherm. Energ. **70**, 20–24 (2011)
96. Freeden, W., Reuter, R.: A constructive method for solving the displacement boundary-value problem of elastostatics by use of global basis systems. Math. Method. Appl. Sci. **12**, 105–128 (1990)
97. Freeden, W., Schreiner, M.: Spherical Functions of Mathematical Geosciences – A Scalar, Vectorial, and Tensorial Setup. Advances in Geophysical and Environmental Mechanics and Mathematics, 1st edn. Springer, Berlin (2009)
98. Frenkel, J.: On the theory of seismic and seismoelectric phenomena in a moist soil. J. Eng. Mech. **131**, 879–887 (2005). Originally published in J. Phys. **3**, 230–241 (1944)
99. Fridleifsson, I.B., Freeston, D.H.: Geothermal energy research and developement. Geothermics **23**, 175–214 (1992)
100. Friedleifsson, I., Ragnarsson, Á.: Survey of Energy Resources, chap. 11. Geothermal Energy, pp. 427–477. World Energy Council, London (2007)
101. Friedman, A.: Partial Differential Equations of Parabolic Type. Prentice-Hall, Englewood Cliffs (1964)
102. Friedrichs, K.O.: On the boundary-value problems of the theory of elasticity and Korn's inequality. Ann. Math. **48**, 441–471 (1947)
103. Gehringer, M., Loksha, V.: Geothermal Handbook: Planning and Financing Power Generation. The International Bank for Reconstruction and Development, The World Bank Group (2012). Energy Sector Management Assistance Program (ESMAP)
104. Gerhards, C.: Spherical decompositions in a global and local framework: theory and an application to geomagnetic modeling. Int. J. Geomath. **1**, 205–256 (2011)
105. Gerhards, C.: Spherical multiscale methods in terms of locally supported wavelets: theory and application to geomagnetic modeling. Ph.D. thesis, University of Kaiserslautern, Geomathematics Group (2011)
106. Gerhards, C.: Locally supported wavelets for the separation of spherical vector fields with respect to their sources. Int. J. Wavelets Multires. Inf. Process. **10**, 1250034:1–1250034:26 (2012)
107. Gerhards, C.: A multiscale power spectrum for the analysis of the lithospheric magnetic field. Int. J. Geomath. **5**, 63–79 (2014)
108. Gerhards, C.: Multiscale modeling of the geomagnetic field and ionospheric currents. In: W. Freeden, Z. Nashed, T. Sonar (eds.) Handbook of Geomathematics, 2nd edn. Springer, New York (2015). Accepted for publication
109. Ghassemi, A.: A review of some rock mechanics issues in geothermal reservoir developement. Geotechn. Geol. Eng. **30**, 647–664 (2012)
110. Ghassemi, A., Cheng, A.H.D., Diek, A., Roegiers, J.C.: A complete plane strain fictitious stress boundary element method for poroelastic media. Eng. Anal. Bound. Elem. **25**, 41–48 (2001)
111. Ghassemi, A., Tarasovs, S.: Three-dimensional modeling of injection induced thermal stresses with an example from Coso. In: Proceedings Twenty-Ninth Workshop on Geothermal Reservoir Engineering, Stanford (2004)
112. Ghassemi, A., Tarasovs, S.: Fracture slip and opening in response to fluid injection into a geothermal reservoir. In: Proceedings Thirty-First Workshop on Geothermal Reservoir Engineering, Stanford (2006)
113. Giusti, E.: Direct Methods in the Calculus of Variations. World Scientific, New Jersey (2003)
114. Glowacka, E., Sarychikhina, O., Nava, F.A.: Subsidence and stress changes in the Cerro Pietro geothermal field, B. C., Mexico. Pure Appl. Geophys. **162**, 2095–2110 (2005)
115. Golberg, M.A.: The method of fundamental solutions for Poisson's equation. Eng. Anal. Bound. Elem. **16**, 205–213 (1995)

116. Golberg, M.A., Chen, C.S.: The method of fundamental solutions for potential, Helmholtz and diffusion problems. In: M.A. Golberg (ed.) Boundary Integral Methods – Numerical and Mathematical Aspects, chap. 4, pp. 103–176. WIT Press, Southampton (1998)
117. Goldemberg, J. (ed.): World energy assesment: energy and the challenge of sustainability. United Nations Development Programme, United Nations Department of Economic and Social Affairs, World Energy Council (2000)
118. Goldemberg, J., Johansson, T.B. (eds.): World energy assesment: overview 2004 update. United Nations Development Programme, United Nations Department of Economic and Social Affairs, World Energy Council (2004)
119. Golub, G.H., Reinsch, C.: Singular value decomposition and least squares solutions. Numer. Math. **14**, 403–420 (1970)
120. Gorzelańczyk, P., Kołodziej, J.A.: Some remarks concerning the shape of the source contour with application of the method of fundamental solutions to elastic torsion of prismatic rods. Eng. Anal. Bound. Elem. **32**, 64–75 (2008)
121. Grothaus, M., Raskop, T.: Limit formulae and jump relations of potential theory in Sobolev spaces. Int. J. Geomath. **1**, 51–100 (2010)
122. Gu, M.H., Fan, C.M., Young, D.L.: The method of fundamental solutions for the multi-dimensional wave equations. J. Mar. Sci. Technol. **19**, 586–595 (2011)
123. Gudmundsson, A., Fjeldskaar, I., Brenner, S.L.: Propagation pathways and fluid transport of hydrofractures in jointed and layered rocks in geothermal fields. J. Volcanol. Geotherm. Res. **116**, 257–278 (2002)
124. Guimaraes, S., Telles, J.C.F.: The method of fundamental solutions for fracture mechanics – Reissner's plate application. Eng. Anal. Bound. Elem. **33**, 1152–1160 (2009)
125. Gunasekera, R.C., Foulger, G.R., Julian, B.R.: Reservoir depletion at The Geysers geothermal area, California, shown by four-dimensional seismic tomography. J. Geophys. Res: Solid Earth **108**, 2156–2202 (2003)
126. Haenel, R., Rybach, L., Stegana, L.: Fundamentals of geothermics. In: R. Haenel, L. Rybach, L. Stegana (eds.) Handbook of Terrestrial Heat-Flow Density Determination, pp. 9–57. Kluwer Academics, Dordrecht (1988)
127. Han, D., Dai, H.H., Qi, L.: Conditions for strong ellipticity of anisotropic elastic materials. J. Elast. **97**, 1–13 (2009)
128. Hazewinkel, M. (ed.): Encyclopedia of Mathematics. Kluwer Academic, Dordrecht (2002)
129. Herrera, I.: Trefftz method: a general theory. Numer. Method. Partial Differ. Equ. **16**, 561–580 (2000)
130. Hicks, T.W., Pine, R.J., Willis-Richards, J., Xu, S., Jupe, A.J., Rodrigues, N.E.V.: A hydro-thermo-mechanical numerical model for HDR geothermal reservoir evaluation. Int. J. Rock Mech. Min. Sci. Geomech. Abstr. **33**, 499–511 (1996)
131. Hon, Y.C., Li, M.: A discrepancy principle for the source points location in using the MFS for solving the BHCP. Int. J. Comput. Method. **6**, 181–197 (2009)
132. Hörmander, L.: The Analysis of Linear Partial Differential Operators I. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, vol. 256. Springer, Berlin (1983)
133. Hu, S.P., Fan, C.M., Chen, C.W., Young, D.L.: Method of fundamental solutions for Stokes' first and second problems. J. Mech. **21**, 25–31 (2005)
134. Ilyasov, M.: A tree algorithm for Helmholtz potential wavelets on non-smooth surfaces: theoretical background and application to seismic data postprocessing. Ph.D. thesis, University of Kaiserslautern, Geomathematics Group (2011)
135. Ingber, M.S., Chen, C.S., Tanski, J.A.: A mesh free approach using radial basis functions and parallel domain decomposition for solving three-dimensional diffusion equations. Int. J. Numer. Method. Eng. **60**, 2183–2201 (2004)
136. Ionescu-Casimir, V.: Problem of linear coupled thermoelasticity. I. Theorems on reciprocity for the dynamic problem of coupled thermoelasticity. Bull. Acad. Pol. Sci., Sér. Sci. Tech. **12**, 473–488 (1964)
137. Jackson, J.D.: Classical Electrodynamics, 3rd edn. Wiley, New York (1998)

138. Jaeger, J.C., Cook, N.G.W., Zimmerman, R.W.: Fundamentals of Rock Mechanics, 4th edn. Blackwell, Malden (2007)
139. Jing, L., Hudson, J.A.: Numerical methods in rock mechanics. Int. J. Rock Mech. Min. Sci. **39**, 409–427 (2002)
140. Johansson, B.T., Lesnic, D.: A method of fundamental solutions for transient heat conduction. Eng. Anal. Bound. Elem. **32**, 697–703 (2008)
141. Johansson, B.T., Lesnic, D.: A method of fundamental solutions for transient heat conduction in layered material. Eng. Anal. Bound. Elem. **33**, 1362–1367 (2009)
142. Johansson, B.T., Lesnic, D., Reeve, T.: A method of fundamental solutions for the one-dimensional inverse Stefan problem. Appl. Math. Model. **35**, 4367–4378 (2011)
143. Johansson, B.T., Lesnic, D., Reeve, T.: A method of fundamental solutions for two-dimensional heat conduction. Int. J. Comput. Math. **88**, 1697–1713 (2011)
144. Johnson, R.: A priori estimates and unique continuation theorems for second order parabolic equations. Trans. Am. Math. Soc. **158**, 167–177 (1971)
145. Johnston, R.L., Fairweather, G.: The method of fundamental solutions for problems in potential flow. Appl. Math. Model. **8**, 265–270 (1984)
146. Jung, R.: Stand und Aussichten der Tiefengeothermie in Deutschland. Erdöl, Erdgas, Kohle **123**, 1–7 (2007)
147. Karageorghis, A., Fairweather, G.: The method of fundamental solutions for the numerical solution of the biharmonic equation. J. Comput. Phys. **69**, 434–459 (1987)
148. Karageorghis, A., Fairweather, G.: The method of fundamental solutions for axisymmetric elasticity problems. Comput. Mech. **25**, 524–532 (2000)
149. Karageorghis, A., Poullikkas, A., Berger, J.R.: Stress intensity factor computation using the method of fundamental solutions. Comput. Mech. **37**, 445–454 (2006)
150. Karageorghis, A., Smyrlis, Y.S., Tsangaris, T.: A matrix decomposition MFS algorithm for certain linear elasticity problems. Numer. Algorithm. **43**, 123–149 (2006)
151. Katsurada, M.: A mathematical study of the charge simulation method II. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **36**, 135–162 (1989)
152. Katsurada, M.: Asymptotic error analysis of the charge simulation method in a Jordan region with analytic boundary. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **37**, 635–657 (1990)
153. Katsurada, M.: Charge simulation method using exterior mapping functions. Jpn. J. Ind. Appl. Math. **11**, 47–61 (1994)
154. Katsurada, M., Okamoto, H.: A mathematical study of the charge simulation method I. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **35**, 507–518 (1988)
155. Katsurada, M., Okamoto, H.: The collocation points of the fundamental solution method for the potential problem. Comput. Math. Appl. **31**, 123–137 (1996)
156. Kaynia, A.M., Banerjee, P.K.: Fundamental solutions of Biot's equations of dynamic poroelasticity. Int. J. Eng. Sci. **31**, 817–830 (1993)
157. Kincaid, D., Cheney, W.: Numerical Analysis – Mathematics of Scientific Computing. Brooks/Cole, Pacific Grove (1991)
158. Kita, E., Kamiya, N.: Trefftz method: an overview. Adv. Eng. Softw. **24**, 3–12 (1995)
159. Kitagawa, T.: On the numerical stability of the method of fundamental solution applied to the Dirichlet problem. Jpn. J. Appl. Math. **5**, 123–133 (1988)
160. Kitagawa, T.: Asymptotic stability of the fundamental solution method. J. Comput. Appl. Math. **38**, 263–269 (1991)
161. Kohl, T., Brenni, R., Eugster, W.: System performance of a deep borehole heat exchanger. Geothermics **31**, 687–708 (2002)
162. Kondapalli, P.S., Shippy, D.J., Fairweather, G.: Analysis of acoustic scattering in fluids and solids by the method of fundamental solutions. J. Acoust. Soc. Am. **91**, 1844–1854 (1992)
163. Kupradze, V.D.: A method for the approximate solution of limiting problems in mathematical physics. USSR Comput. Math. Math. Phys. **4**, 199–205 (1964)
164. Kupradze, V.D.: Potential Methods in the Theory of Elasticity, Translated from the Russian by H. Gutfreund, Translation edited by I. Meroz. Israel Program for Scientific Translations, Jerusalem (1965)

165. Kupradze, V.D.: On the approximate solution of problems in mathematical physics. Russ. Math. Surv. **22**, 58–108 (1967)
166. Kupradze, V.D., Aleksidze, M.A.: An approximate method of solving certain boundary-value problems. Soobŝ. Akad. Nauk Gruz. SSR (Bull. Acad. Sci. Georgian SSR) **30**, 529–536 (1963). In Russian
167. Kupradze, V.D., Aleksidze, M.A.: The method of functional equations for the approximate solution of certain boundary value problems. USSR Comput. Math. Math. Phys. **4**, 82–126 (1964)
168. Kurashige, M., Sato, K., Imai, K.: Mandel and Cryer problems of fluid-saturated foam with negative Poisson's ratio. Acta Mech. **175**, 25–43 (2005)
169. Lai, M., Krempl, E., Ruben, D.: Introduction to Continuum Mechanics, 4th edn. Butterworth-Heinemann, Oxford (2010)
170. Landau, L.D., Pitaevskii, L.P., Lifshitz, E.M., Kosevich, A.M.: Theory of Elasticity. Theoretical Physics, vol. 7, 3rd edn. Butterworth-Heinemann, Oxford (1986)
171. LeVeque, R.J.: Lectures in mathematics, ETH Zürich. In: O.E. Lanford (ed.) Numerical Methods for Conversation Laws, 2nd edn. Birkhäuser, Basel (1992)
172. Li, X.: On convergence of the method of fundamental solutions for solving the Dirichlet problem of Poisson's equation. Adv. Comput. Math. **23**, 265–277 (2005)
173. Li, X.: Convergence of the method of fundamental solutions for Poisson's equation on the unit sphere. Adv. Comput. Math. **28**, 269–282 (2008)
174. Li, X.: Rate of convergence of the method of fundamental solutions and hyperinterpolation for modified Helmholtz equations on the unit ball. Adv. Comput. Math. **29**, 393–413 (2008)
175. Lions, J.L.: Equations Differentielles Operationelles et Problèmes aux Limites. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, vol. 111. Springer, Berlin (1961)
176. Liu, G., Xie, K., Zheng, R.: Model of nonlinear coupled thermo-hydro-elastodynamics response for a saturated poroelastic medium. Sci. China, Ser. E: Technol. Sci. **52**, 2373–2383 (2009)
177. Lopatnikov, S.L., Gillespie, J.W., Jr.: Poroelasticity-I: governing equations of the mechanics of fluid-saturated porous material. Transp. Porous Media **84**, 471–492 (2010)
178. Lopatnikov, S.L., Gillespie, J.W., Jr.: Poroelasticity-II: on the equilibrium state of the fluid-filled penetrable poroelastic body. Transp. Porous Media **89**, 475–486 (2011)
179. Lopatnikov, S.L., Gillespie, J.W., Jr.: Poroelasticity-III: conditions on the interface. Transp. Porous Media **93**, 597–607 (2012)
180. Lowitzsch, S.: Error estimates for matrix-valued radial basis function interpolation. J. Approx. Theory **137**, 238–249 (2005)
181. Lowitzsch, S.: Matrix-valued radial basis functions: stability estimates and applications. Adv. Comput. Math. **23**, 299–315 (2005)
182. Luchko, Y., Punzi, A.: Modeling anomalous heat transport in geothermal reservoirs via fractional diffusion equations. Int. J. Geomath. **1**, 257–276 (2011)
183. Lund, J.W.: Direct utilization of geothermal energy. Energies **3**, 1443–1471 (2010)
184. Malgrange, B.: Éxistence et Approximation des Solutions des Équations aus Dérivées Partielles et des Équations de Convolution. Ann. Inst. Fourier **6**, 271–355 (1956)
185. Mandel, J.: Consolidation de Sols (Étude Mathématique). Géotechnique **3**, 287–299 (1953)
186. Manolis, G.D., Beskos, D.E.: Integral formulation and fundamental solutions of dynamic poroelasticity and thermoelasticity. Acta Mech. **76**, 89–104 (1989)
187. Marin, L.: An alternating iterative MFS algorithm for the Cauchy problem for the modified Helmholtz equation. Comput. Mech. **45**, 665–677 (2010)
188. Marsden, J.E., Hughes, T.J.R.: Mathematical Foundations of Elasticity. Dover, Mineola (1994)
189. Mathon, R., Johnston, R.L.: The approximate solution of elliptic boundary-value problems by fundamental solutions. SIAM J. Numer. Anal. **14**, 638–650 (1977)
190. Matlab: version R2013a. The MathWorks (2013)

191. Mayer, C.: A wavelet approach to the Stokes problem. Habilitation Thesis, University of Kaiserslautern, Geomathematics Group (2007)
192. Mayer, C., Freeden, W.: Stokes problem, layer potentials and regularizations, multiscale applications. In: W. Freeden, Z. Nashed, T. Sonar (eds.) Handbook of Geomathematics, 2nd edn. Springer, New York (2015). Accepted for publication
193. Mazzucato, A.L., Nistor, V.: Well-posedness and regularity for the elasticity equation with mixed boundary conditions on polyhedral domains and domains with cracks. Arch. Ration. Mech. Anal. **195**, 25–73 (2010)
194. McLean, W.: Strongly Elliptic Systems and Boundary Integral Equations. Cambridge University Press, Cambridge (2000)
195. Cirne de Mederios, G., Partridge, P.W.: The method of fundamental solutions with dual reciprocity for thermoelasticity. In: International Workshop on Meshfree Methods, Lisbon (2003)
196. Cirne de Mederios, G., Partridge, P.W., Bandão, J.O.: The method of fundamental solutions with dual reciprocity for some problems in elasticity. Eng. Anal. Bound. Elem. **28**, 453–461 (2004)
197. Menéndez, C., Nieto, P.J.G., Ortega, F.A., Bello, A.: Mathematical modelling and study of the consolidation of an elastic saturated soil with an incompressible fluid by FEM. Math. Comput. Model. **49**, 2002–2018 (2009)
198. Miehe, C., Méndez Diez, J., Göktepe, S., Schänzel, L.M.: Coupled thermoviscoplasticity of glassy polymers in the logarithmic strain space based on the free volume theory. Int. J. Solid. Struct. **48**, 1799–1817 (2010)
199. Mikhailov, V.P.: Partial Differential Equations. MIR Publishers, Moscow (1978)
200. Mikhlin, S.G.: Mathematical Physics, an Advanced Course. North-Holland, Amsterdam (1970)
201. Möhringer, S.: Decorrelation of gravimetric data. Ph.D. thesis, University of Kaiserslautern, Geomathematics Group (2014)
202. Moldovan, I.D., Teixera de Freitas, J.A.: Hybrid-trefftz stress and displacement elements for dynamic analysis of bounded and unbounded saturated porous media. Comput. Assist. Mech. Eng. Sci. **15**, 289–303 (2008)
203. Moler, C.: Cleve's corner: professsor SVD. The MathWorks News & Notes (2006). Downloaded from http://www.mathworks.de/company/newsletters/articles/professor-svd.html. On 28 Feb 2014
204. Mosconi, M.: A variational approach to porous elastic bodies. Z. Angew. Math. Phys. ZAMP **56**, 548–558 (2005)
205. Mossop, A., Segall, P.: Subsidence at The Geysers geothermal field, N. California from a comparison of GPS and leveling surveys. Geophys. Res. Lett. **24**, 1839–1842 (1997)
206. Müller, C., Kersten, H.: Zwei Klassen vollständiger Funktionensysteme zur Behandlung der Randwertaufgaben der Schwingungsgleichung $\Delta U + k^2 U = 0$. Math. Method. Appl. Sci. **2**, 48–67 (1980)
207. Nakao, S., Ishido, T.: Pressure-transient behavior during cold water injection into geothermal wells. Geothermics **27**, 401–413 (1998)
208. Nardini, D., Brebbia, C.A.: A new approach for free vibration analysis using boundary elements. In: C.A. Brebbia (ed.) Boundary Element Methods in Engineering: Proceedings of the Fourth International Seminar, Southhampton, Sept 1982, pp. 312–326. Springer, Berlin (1982)
209. Naumovich, A.: Efficient numerical methods for the Biot poroelasticity system in multilayered domains. Ph.D. thesis, University of Kaiserslautern, Department of Mathematics (2007)
210. Navarro, C.B., Quintanilla, R.: On existence and uniqueness in incremental thermoelasticity. J. Appl. Math. Phys. **35**, 206–215 (1984)
211. Nennig, P., Perrey-Debain, E., Chazot, J.D.: The method of fundamental solutions for acoustic wave scattering by a single and a periodic array of poroelastic scatterers. Eng. Anal. Bound. Elem. **35**, 1019–1028 (2011)

212. Neumaier, A.: Solving ill-conditioned and singular linear systems: a tutorial on regularization. SIAM Rev. **40**, 636–666 (1998)
213. Nowacki, W.: Thermoelasticity. Aeronautics and Astronautics. Pergamon Press, Oxford (1986)
214. Ostermann, I.: A mathematical model for 3D heat transport in hydrothermal systems. World Min. Surf. Undergr. **63**, 280–289 (2011)
215. Ostermann, I.: Modeling heat transport in deep geothermal systems by radial basis functions. Ph.D. thesis, University of Kaiserslautern, Geomathematics Group (2011)
216. Ostermann, I.: Three-dimensional modeling of heat transport in deep hydrothermal reservoirs. Int. J. Geomath. **2**, 37–68 (2011)
217. Owczarek, S.: A Galerkin method for Biot consolidation model. Math. Mech. Solid. **15**, 42–56 (2010)
218. Pan, E.: Green's functions in layered poro-elastic half-spaces. Int. J. Numer. Anal. Method. Geomech. **23**, 1631–1653 (1999)
219. Partridge, P.W., Sensale, B.: The method of fundamental solutions with dual reciprocity for diffusion and diffusion-convection using subdomains. Eng. Anal. Bound. Elem. **24**, 633–641 (2000)
220. Paschen, H., Oertel, D., Grünwald, R.: Möglichkeiten geothermischer Stromerzeugung in Deutschland. TAB Arbeitsbericht 84, Deutscher Bundestag, Ausschuss für Bildung, Forschung und Technikfolgeabschätzung (2003)
221. Phillips, P.J.: Finite element method in linear poroelasticity: theoretical and computational results. Ph.D. thesis, University of Texas, Austin (2005)
222. Phillips, P.J., Wheeler, M.F.: A coupling of mixed and continuous Galerkin finite-element methods for poroelasticity I: the continuous in time case. Comput. Geosci. **11**, 131–144 (2007)
223. Phillips, P.J., Wheeler, M.F.: A coupling of mixed and continuous Galerkin finite-element methods for poroelasticity II: the discrete-in-time case. Comput. Geosci. **11**, 145–158 (2007)
224. Phillips, P.J., Wheeler, M.F.: A coupling of mixed and discontinuous Galerkin finite-element methods for poroelasticity. Comput. Geosci. **12**, 417–435 (2008)
225. Phillips, P.J., Wheeler, M.F.: Overcoming the problem of locking in linear elasticity and poroelasticity: an heuristic approach. Comput. Geosci. **13**, 5–12 (2009)
226. Poullikkas, A., Karageorghis, A., Georgiou, G.: The numerical solution of three-dimensional Signorini problems with the method of fundamental solutions. Eng. Anal. Bound. Elem. **25**, 221–227 (2001)
227. Predeleanu, M.: Reciprocal theorem in the consolidation theory of porous media. An. Univ. Bucur. Ser. Ştiinţ. Nat. Mat.-Mec. **17**, 75–79 (1968)
228. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Fortran Numerical Recipes, 2nd edn. Cambridge University Press, Cambridge (2001)
229. Prudnikov, A.P., Brychkov, Y.A., Marichev, O.I.: Direct Laplace Transforms. Integrals and Series, vol. 4. Gordeon and Breach Science Publishers, Amsterdam (1992)
230. Prudnikov, A.P., Brychkov, Y.A., Marichev, O.I.: Inverse Laplace Transforms. Integrals and Series, vol. 5. Gordeon and Breach Science Publishers, Amsterdam (1992)
231. Quintanilla, R.: Spatial stability for the quasi-static problem of thermoelasticity. J. Elast. **76**, 93–105 (2004)
232. Raskop, T.: The analysis of oblique boundary problems and limit formulae motivated by problems from geomathematics. Ph.D. thesis, University of Kaiserslautern, Functional Analysis and Stochastic Analysis Group (2009)
233. Redekop, D.: Fundamental solutions for the collation [sic] method in planar elastostatics. Appl. Math. Model. **6**, 390–393 (1982)
234. Renardy, M., Rogers, R.C.: An Introduction to Partial Differential Equations. Texts in Applied Mathematics. Springer, New York (1993)
235. Reutskiy, S.Y.: The method of approximate fundamental solutions (MAFS) for Stefan problems. Eng. Anal. Bound. Elem. **36**, 281–292 (2012)

236. Ritz, W.: Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. J. rein. angew. Math. **135**, 1–6 (1909)
237. Rose, N.J. (ed.): Mathematical Maxims and Minims. Rome Press, Raleigh (1988). Illustrated by J. de Pillis
238. Rudin, W.: Functional Analysis. McGraw-Hill, New York (1973)
239. Runge, C.: Zur Theorie der eindeutigen analytischen Functionen. Acta Math. **6**, 229–234 (1885)
240. Rutqvist, J., Stephansson, O.: The role of hydromechanical coupling in fractured rock engineering. Hydrogeol. J. **11**, 7–40 (2003)
241. Saemundsson, K., Axelsson, G., Steingrimsson, B.: Geothermal systems in global perspectives. In: Short Course V on Conceptual Modelling of Geothermal Systems. United Nations University Geothermal Training Programme, La Geo, Santa Tecla, El Salvador (2013)
242. Saint-Macary, P.: Analyse Mathématique de Modèles de Diffusion en Milieu Poreux élastique. Ph.D. thesis, L'Université de Pau et des Pays de l'Adour (2004)
243. Sanyal, S.K.: Classification of geothermal systems – a possible scheme. In: Proceedings of the Thirtieth Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford (2005)
244. Saut, J.C., Scheurer, B.: Unique continuation for some evolution equations. J. Differ. Equ. **66**, 118–139 (1987)
245. Sauter, M., Wieners, C.: Robust estimates for the approximation of the dynamic consolidation problem. IMA J. Numer. Anal. **30**, 832–856 (2010)
246. Schaback, R.: Adaptive numerical solution of MFS systems. In: C.S. Chen, A. Karageorghis, Y.S. Smyrlis (eds.) The Method of Fundamental Solutions – A Meshless Method, pp. 1–27. Dynamic Publishers, Atlanta (2008)
247. Schanz, M.: Application of 3D time domain boundary element formulation to wave propagation in poroelastic solids. Eng. Anal. Bound. Elem. **25**, 363–376 (2001)
248. Schanz, M., Pryl, D.: Dynamic fundamental solutions for compressible and incompressible modeled poroelastic continua. Int. J. Solid. Struct. **41**, 4047–4073 (2004)
249. Schellschmidt, R., Sanner, B., Pester, S., Schulz, R.: Geothermal energy use in Germany. In: Proceedings World Geothermal Congress, Bali (2010)
250. Schulz, R.: Aufbau eines geothermischen Informationssystems für Deutschland. Tech. Rep., Leibniz-Institut für Angewandte Geophysik, Hannover (2009)
251. Selvadurai, A.P.S.: The analytical method in geomechanics. Appl. Mech. Rev. **60**, 87–106 (2007)
252. Sensale-Rodriguez, B., Sensale, B., Leitão, V.M.A., Peixeiro, C.: Microstrip antenna analysis using the method of fundamental solutions. Int. J. Numer. Model.: Electron. Net. Device. Field. **21**, 563–581 (2008)
253. Showalter, R.E.: Monotone Operators in Banach Space and Nonlinear Partial Differential Equations. Mathematical Surveys and Monographs, vol. 49. American Mathematical Society, Providence (1996)
254. Showalter, R.E.: Diffusion in poro-elastic media. J. Math. Anal. Appl. **251**, 310–340 (2000)
255. Showalter, R.E.: Diffusion in deformable media. In: J. Chadam, A. Cunningham, R.E. Ewing, P. Ortoleva, M.F. Wheeler (eds.) Resource Recovery, Confinement and Remedation of Environmental Hazards. The IMA Volumes in Mathematics and its Applications, vol. 131. Springer, New York (2002)
256. Showalter, R.E., Momken, B.: Single-phase flow in composite poroelastic media. Math. Method. Appl. Sci. **25**, 115–139 (2002)
257. Showalter, R.E., Stefanelli, U.: Diffusion in poro-plastic media. Math. Method. Appl. Sci. **27**, 2131–2151 (2004)
258. Showalter, R.E., Su, N.: Partially saturated flow in a poro-elastic medium. Discret. Contin. Dyn. Syst. Ser. B **1**, 403–420 (2001)
259. Showalter, R.E., Su, N.: Partially saturated flow in a composite poroelastic medium. In: J.L. Auriault, C. Geindreau, P. Royer, J.F. Bloch (eds.) Poromechanics II, Proceedings of the Second Biot Conference on Poromechanics, pp. 549–554. Balkema, Grenoble (2002)

260. Skopetskii, V.V., Marchenko, O.A.: Formulation and analysis of problems for dynamic systems of inhomogeneous two-phase media. Cybern. Syst. Anal. **41**, 865–878 (2005)
261. Smyrlis, Y.S.: The method of fundamental solutions: a weighted least-squares approach. BIT Numer. Math. **46**, 163–194 (2006)
262. Smyrlis, Y.S.: Applicability and applications of the method of fundamental solutions. Math. Comput. **78**, 1399–1434 (2009)
263. Smyrlis, Y.S.: Density results with linear combinations of translates of fundamental solutions. J. Approx. Theory **161**, 617–633 (2009)
264. Smyrlis, Y.S.: Mathematical foundation of the MFS for certain elliptic systems in linear elasticity. Numer. Math. **112**, 319–340 (2009)
265. Smyrlis, Y.S., Karageorghis, A.: Numerical analysis of the MFS for certain harmonic problems. ESAIM: Math. Model. Numer. Anal. **38**, 495–517 (2004)
266. Smyrlis, Y.S., Karageorghis, A.: Efficient implementation of the MFS: the three scenarios. J. Comput. Appl. Math. **227**, 83–92 (2009)
267. Smyrlis, Y.S., Karageorghis, A.: The under-determined version of the MFS: taking more sources than collocation points. Appl. Numer. Math. **60**, 337–357 (2010)
268. Stoer, J., Bulirsch, R.: Introduction to Numerical Analysis, 2nd edn. Springer, New York (1993). Translated from German by R. Bartels, W. Gautschi, and C. Witzgall
269. Strang, G.: Linear Algebra and Its Applications. Brooks/Cole, Toronto (1988)
270. Stynes, M.: Steady-state convection-diffusion problems. Acta Numer. **14**, 445–508 (2005)
271. von Terzaghi, K.: Die Berechnung der Durchlässigkeitsziffer des Tones aus dem Verlauf der hydrodynamischen Spannungserscheinungen. Sitz.-ber. Akad. Wiss., Wien, Math.-Nat.-wiss. Kl., Abt. IIa **132**, 105–124 (1923)
272. von Terzaghi, K.: Theoretical Soil Mechanics. Wiley, New York (1943)
273. Tester, J.W., Anderson, B.J., Batchelor, A.S., Blackwell, D.D., DiPippo, R., Drake, E.M., Garnish, J., Livesay, B., Moore, M.C., Nichols, K., Petty, S., Toksösz, M.N., Veatch Jr., R.W., Baria, R., Augustine, C., Murphy, E., Negraru, P., Richards, M.: The Future of geothermal energy. Tech. Rep., Idaho National Laboratory, copyright by Massachussetts Institute of Technology (2006)
274. Trefftz, E.: Ein Gegenstück zum Ritzschen Verfahren. In: Proceedings of the Second International Congress on Applied Mechanics, Zürich (1926)
275. Tsai, C.C.: Solutions of slow Brinkman flows using the method of fundamental solutions. Int. J. Numer. Method. Fluid. **56**, 927–940 (2008)
276. Tsai, C.C.: The method of fundamental solutions with dual reciprocity for three-dimensional thermoelasticity under arbitrary body forces. Eng. Comput. Int. J. Comput. Eng. Softw. **26**, 229–244 (2009)
277. Tsai, C.C., Hsu, T.W.: The method of fundamental solutions for oscillatory and porous buoyant flows. Comput. Fluid. **39**, 696–708 (2010)
278. Tsai, C.C., Hsu, T.W.: A meshless numerical method for solving slow mixed convections in containers with discontinuous boundary data. Int. J. Numer. Method. Fluid. **66**, 377–402 (2011)
279. Tsai, C.C., Young, D.L., Fan, C.M., Chen, C.W.: MFS with time-dependent fundamental solutions for unsteady Stokes equations. Eng. Anal. Bound. Elem. **30**, 897–908 (2006)
280. Tsangaris, T., Smyrlis, Y.S., Karageorghis, A.: Numerical analysis of the method of fundamental solutions for harmonic problems in annular domains. Numer. Method. Partial Differ. Equ. **21**, 507–539 (2005)
281. Valtchev, S.S., Roberty, N.C.: A time-marching MFS scheme for heat conduction problems. Eng. Anal. Bound. Elem. **32**, 480–493 (2008)
282. Vgenopoulou, I., Beskos, D.E.: Dynamic behavior of saturated poroviscoelastic media. Acta Mech. **95**, 185–195 (1992)
283. Visual Numerics (ed.): IMSL User Manual. Rogue Wave Software, Boulder, Colorado (2010)
284. Walsh, J.L.: The approximation of harmonic functions by harmonic polynomials and by harmonic rational functions. Bull. Am. Math. Soc. **35**, 499–544 (1929)

285. Wang, H.F.: Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrogeology. Princeton University Press, Princeton (2000)
286. Wang, W., Kosakowski, G., Kolditz, O.: A parallel finite element scheme for thermo-hydro-mechanical (THM) coupled problems in porous media. Comput. Geosci. **35**, 1631–1641 (2009)
287. Weinstock, B.M.: Uniform approximation by solutions of elliptic equations. Proc. Am. Math. Soc. **41**, 513–517 (1973)
288. Wen, P.H., Liu, Y.W.: The fundamental solution of poroelastic plate saturated by fluid and its applications. Int. J. Numer. Anal. Method. Geomech. **34**, 689–709 (2010)
289. Widder, D.: The Laplace Transform. Princeton University Press, Princeton (1941)
290. Wiebe, T., Antes, H.: A time domain integral formulation of dynamic poroelasticity. Acta Mech. **90**, 125–137 (1991)
291. Wirth, B., Sobey, I.: Analytic solution during an infusion test of the linear unsteady poroelastic equations in a spherically symmetric model of the brain. Math. Med. Biol. **26**, 25–61 (2009)
292. Wloka, J.: Partial Differential Equations. Cambridge University Press, Cambridge (1992)
293. Yin, S.: Geomechanics-reservoir modeling by displacement discontinuity-finite element method. Ph.D. thesis, University of Waterloo (2008)
294. Yin, S., Rothenburg, L., Dusseault, M.B.: 3D coupled displacement discontinuity and finite element analysis of reservoir behavior during production in semi-infinite domain. Transp. Porous Media **65**, 425–441 (2006)
295. Yosida, K.: Functional Analysis. Springer, Berlin (1980)
296. Young, D.L., Chen, C.H., Fan, C.M., Shen, L.H.: The method of fundamental solutions with eigenfunctions expansion method for 3d nonhomogeneous diffusion equations. Numer. Method. Partial Differ. Equ. **25**, 195–211 (2009)
297. Young, D.L., Chen, C.W., Fan, C.M., Tsai, C.C.: The method of fundamental solutions with eigenfunction expansion method for nonhomogeneous diffusion equation. Numer. Method. Partial Differ. Equ. **22**, 1173–1196 (2006)
298. Young, D.L., Fan, C.M., Hu, S.P., Atluri, S.N.: The Eulerian-Lagrangian method of fundamental solutions for two-dimensional unsteady Burgers' equations. Eng. Anal. Bound. Elem. **32**, 395–412 (2008)
299. Young, D.L., Jane, S.J., Fan, C.M., Murugesan, K., Tsai, C.C.: The method of fundamental solutions for 2D and 3D Stokes problems. J. Comput. Phys. **211**, 1–8 (2006)
300. Young, D.L., Tsai, C.C., Murugesan, K., Fan, C.M., Chen, C.W.: Time-dependent fundamental solutions for homogeneous diffusion problems. Eng. Anal. Bound. Elem. **28**, 1463–1473 (2004)
301. Ženíšek, A.: The existence and uniqueness theorem in Biot's consolidation theory. Appl. Math. **29**, 194–211 (1984)
302. Zhou, H., Pozrikidis, C.: Adaptive singularity method for Stokes flow past particles. J. Comput. Phys. **117**, 79–89 (1995)
303. Zhou, X.X., Ghassemi, A.: Three-dimensional poroelastic simulation of hydraulic and natural fractures using the displacement discontinuity method. In: Proceedings Thirty-Fourth Workshop on Geothermal Reservoir Engineering, Stanford (2009)
304. Zubkov, V.V., Koshelev, V.F., Lin'kov, A.M.: Numerical modeling of hydraulic fracture initiation and developement. J. Min. Sci. **43**, 40–56 (2007)