

**Adaptive Multiscale Finite Element  
Method for Subsurface Flow Simulation**



# Adaptive Multiscale Finite Element Method for Subsurface Flow Simulation

## PROEFSCHRIFT

ter verkrijging van de graad van doctor  
aan de Technische Universiteit Delft,  
op gezag van de Rector Magnificus, prof.ir. K.C.A.M. Luyben,  
voorzitter van het College voor Promoties,  
in het openbaar te verdedigen op maandag 15 november 2010 om 15:00 uur

door Johannes Maria VAN ESCH

civil ingenieur  
geboren te Heemstede

Dit proefschrift is goedgekeurd door de promotor:  
Prof.dr.ir. F.B.J. Barends

Samenstelling promotie-commissie:

Rector Magnificus,	Technische Universiteit Delft, voorzitter
Prof.dr.ir. F.B.J. Barends,	Technische Universiteit Delft, promotor
Prof.dr. J. Bruining,	Technische Universiteit Delft
Prof.dr. M.A. Hicks,	Technische Universiteit Delft
Prof.dr.ir. T.N. Olsthoorn,	Technische Universiteit Delft
Prof.dr. R.J. Schotting,	Universiteit Utrecht
Prof.dr.-ing. P.A. Vermeer,	Universität Stuttgart
Dr.ir. J.A.M. Teunissen,	Deltares

*Printed by:*

Wöhrmann Print Service  
P.O. Box 92  
7202 CZ Zutphen  
The Netherlands  
Telephone: +31 575 58 53 00  
E-mail: wps@wps.nl

ISBN 978-90-8570-608-3

© 2010 by J.M. van Esch

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written consent of the publisher.

Printed in The Netherlands

# Summary

Natural geological formations generally show multiscale structural and functional heterogeneity evolving over many orders of magnitude in space and time. In subsurface hydrological simulations the geological model focuses on the structural hierarchy of physical sub units and the flow model addresses the functional hierarchy of the flow process. Flow quantities like pressure, flux and dissipation relate to each other by constitutive relations and structural sub-unit parameters like porosity and hydraulic permeability. Hydraulic permeability includes the (steady state) intrinsic permeability of the solid phase and the (time dependent) relative permeability. The permeability is the dominant parameter and a highly heterogeneous parameter affecting the groundwater flow at field scale.

Laboratory experiments provide measurements of the sub-unit parameters on a fine scale. If laboratory measurements are treated stochastically within the geological model, then the structural model of the subsurface should be built on the same scale as these indicating measurements. Fully resolved flow simulations on the field scale are however intractable and a new adaptive multiscale technique has therefore been developed. Though constitutive relations may change at different scales, Darcy's law is supposed to remain valid on both laboratory scale and field scale.

Nowadays upscaling methods are applied, which aim to propagate information over this hierarchy of scales both functionally and structurally from the fine scale to the coarse scale and not vice versa, by computing effective or equivalent material behavior. Permeability is not an additive parameter so it is not possible to calculate an equivalent coarse scale permeability as a simple average of fine scale measurements. A flow criterion or a criterion about energy dissipation often defines an equivalent permeability. Only if the scale of variation is much smaller than the coarse observation scale then the equivalent permeability matches the effective permeability. This effective permeability is a constant second order tensor variable on the coarse scale, whereas the equivalent permeability is non-unique and depends on the boundary conditions of the sample domain. The effective permeability holds for discrete hierarchical systems where scales can be decomposed.

The newly developed adaptive multiscale technique extends the original two-level multiscale finite element method to a hierarchy of scales. Multiscale finite element methods capture the fine scale behavior on a coarse mesh by multiscale basis functions. The weights of the multiscale basis functions follow from local flow simulation. The procedure removes fine scale nodes from the subdomain, but introduces errors at the subdomain boundaries. The two-level method forms a class of subdomain decom-

position techniques. It can be shown that a sequential implementation of the method is not faster than an optimal solver like the full multigrid solver, however the method is suitable for parallel implementation.

The proposed method is based on a conformal nodal finite element formulation over simplex elements. The conformal finite element method obtains mass conservation on a nodal basis, and does not preserve continuity of flux over the inter-element boundaries. A mesh refinement criterion detects zones in which large errors occur over the element edges, and an adaptive refinement procedure enriches the mesh locally to correct the error in the velocity field. Multiscale basis functions follow from solving local flow problems over patches of simplex elements. Linear boundary conditions close over-sampled subdomains and reduce the effect of the imposed boundary condition on the patch. This procedure obtains more accurate coarse scale behavior than a procedure that operates on the patch directly. However, over-sampled local flow problems introduce discontinuities in the basis functions and introduce new nodal connectivities on the coarser scale. For this reason closure of the local flow problems by dimensionally reduced flow problems is preferred. A second refinement criterion compares oversampled and non-oversampled function values. Pressure-dissipation averaging approximates the multiscale coarse grid operator and supports a functional adaptive formulation. The multiscale averaging procedure computes equivalent permeability tensor components, and reproduces the sparse matrix structure on the coarse scale. The loading cases for the local problems follow from a summation of multiscale basis functions. The multiscale basis functions extrapolate the coarse scale solution to the fine scale. On this scale discontinuities in the velocity field are detected and compared to the refinement criterion. The computed equivalent permeability is used in the framework of a geometric multigrid solver to compute the coarse scale operators. The hierarchy of multiscale basis functions, which relate the pressure on each coarse level to the next fine level, generates the intergrid transfer operators.

The proposed approach provides a robust and efficient algorithm, based on the concept of the multiscale finite element method, for simulating partly saturated subsurface flow and fully coupled solute transport and heat transport through hierarchical heterogeneous formations. The multiscale finite element formulation produces numerically homogenized discrete flow equations, and upscales the permeability. The adaptive formulation obtains locally refined velocity fields, which support accurate transport computations. The performance of the method is illustrated by a set of realistic case studies.

John van Esch.

# Contents

<b>Summary</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Mathematical model</b>	<b>9</b>
2.1 Balance equations . . . . .	10
2.1.1 Mass balance equation . . . . .	10
2.1.2 Momentum balance equation . . . . .	12
2.1.3 Energy balance equation . . . . .	15
2.2 Constitutive relations . . . . .	17
2.2.1 Porosity . . . . .	18
2.2.2 Saturation . . . . .	20
2.2.3 Permeability . . . . .	22
2.2.4 Density . . . . .	23
2.2.5 Viscosity . . . . .	24
2.3 Problem definition . . . . .	24
2.3.1 System of differential equations . . . . .	24
2.3.2 Initial conditions . . . . .	28
2.3.3 Boundary conditions . . . . .	28
<b>3 Numerical model</b>	<b>33</b>
3.1 Basic equations . . . . .	34
3.1.1 Flow equation . . . . .	34
3.1.2 Solute transport equation . . . . .	40
3.1.3 Heat transport equation . . . . .	45
3.2 Problem definition . . . . .	49
3.2.1 System of algebraic equations . . . . .	50
3.2.2 Initial conditions . . . . .	58
3.2.3 Boundary conditions . . . . .	59
<b>4 Adaptive multiscale method</b>	<b>61</b>
4.1 Adaptive formulation . . . . .	62
4.1.1 Discrete adaptivity . . . . .	62
4.1.2 Functional adaptivity . . . . .	70

---

4.2	Multiscale formulation . . . . .	74
4.2.1	Multiscale basis functions . . . . .	75
4.2.2	Coarse mesh equations . . . . .	81
4.3	Multigrid formulation . . . . .	82
<b>5</b>	<b>Model verification</b>	<b>89</b>
5.1	Saturated and unsaturated flow . . . . .	90
5.1.1	Lamb and Whitman's problem . . . . .	90
5.1.2	Forsyth et al's problem . . . . .	94
5.2	Solute transport and heat transport . . . . .	98
5.2.1	Henry's problem . . . . .	99
5.2.2	Elder's heat problem . . . . .	101
<b>6</b>	<b>Flow simulation</b>	<b>107</b>
6.1	Marine system . . . . .	108
6.1.1	Geological model . . . . .	108
6.1.2	Flow model . . . . .	109
6.2	Fluvial system . . . . .	114
6.2.1	Geological model . . . . .	115
6.2.2	Flow model . . . . .	117
<b>7</b>	<b>Conclusions</b>	<b>121</b>
<b>A</b>	<b>Continuum mechanics</b>	<b>123</b>
<b>B</b>	<b>Finite element method</b>	<b>137</b>
<b>C</b>	<b>Multigrid method</b>	<b>153</b>
	<b>Notations</b>	<b>173</b>
	<b>Bibliography</b>	<b>177</b>
	<b>Samenvatting</b>	<b>187</b>

# Chapter 1

## Introduction

Natural earth materials are of irregular and complex nature [55, 31]. In general soils are not homogeneous, nor uniformly random but may contain multiple, nested, natural length and time scales or even continuous evolving scales [29]. The multiscale heterogeneity evolves over many orders of magnitude in space and time. As for many multiscale problems two types of heterogeneity are often considered: structural heterogeneity and functional heterogeneity. Structural heterogeneity deals with the hierarchy of physical sub units and functional heterogeneity deals with the hierarchy of processes. Both structural and functional hierarchies are mostly strongly connected and may be discrete or continuous. The discrete type incorporates a finite number of nested structural subunits or subprocesses. The number of subunits and subprocesses goes to infinity without a clear cut decomposition for the continuous type. Periodic porous media for instance show a discrete hierarchy and fractal porous media show a continuous hierarchy.

In subsurface hydrological simulations, the geological model focuses on the structural hierarchy of physical subunits and the flow model addresses the functional hierarchy of the flow process. Flow quantities like pressure, flux and dissipation relate to each other by constitutive relations and structural subunit parameters like porosity and hydraulic permeability. Porosity generally varies over a single order of magnitude. The hydraulic permeability however may vary over eight orders of magnitude due to variations in soil structure and soil type or in the degree of saturation. The scale of variability of the permeability may be a fraction of a meter, and time variations may take place over a period of an hour. Figure 1.1 presents two single scale periodic permeability fields. Both single scale fields contribute (by adding the logarithmic values attached to their subunits) to the multiscale periodic field shown in the most right picture.

In heterogeneous fields the window of observation  $l^d$  and the resolution of the measurement  $(l/n)^d$  may alter the measured value of the field variable. Here,  $l$  is length,  $d$  corresponds to the dimension and  $n$  denotes the number of pixels in one dimension. An apparatus may for instance consider a pixel to be totally void if the void space intersects the pixel area, and totally solid if not. Under these assumptions the porosity is the weighted sum of the pixel areas labeled void divided by the observation area.

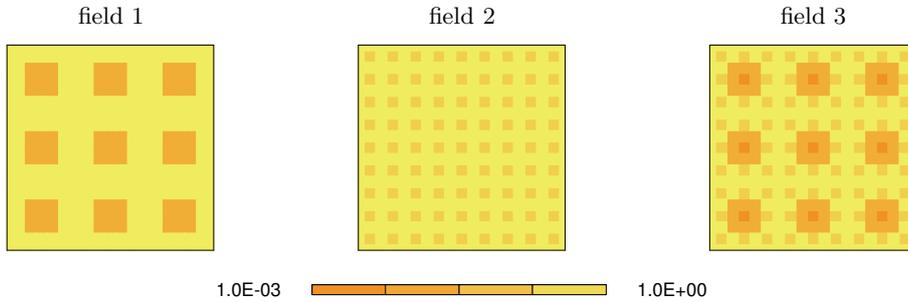


Figure 1.1: Periodic permeability fields.

The void space of the porous medium is fractal if the porosity goes to zero as the number of pixels  $n^d$  goes to infinity. If the porosity goes to one then the solid space is generally fractal. Figure 1.2 presents a fractal field. The resolution has been increased from the left to the right (low porosity values prevail). If not fractal then the porosity

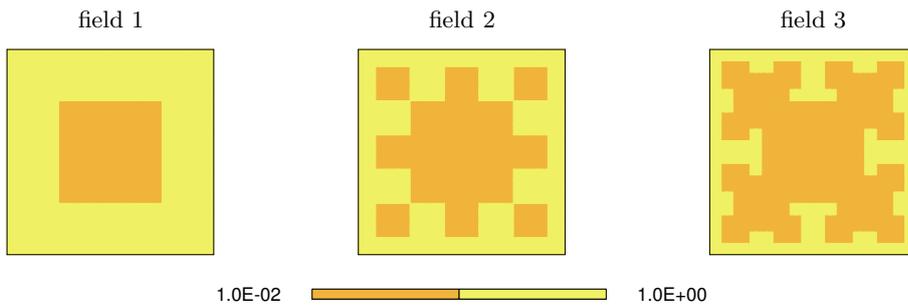


Figure 1.2: Fractal porosity fields.

goes monotonically to some constant average if the number of observation pixels goes to infinity. For a common focal point and a fixed resolution the window of observation may also be changed. The porosity as a function of observation area may show one or more plateaus, if it possesses a discrete hierarchy and the porosity has one or more natural scales. These asymptotic limits correspond to representative elementary volumes (REV). If there are no plateaus then the heterogeneity is said to evolve with respect to the observation area of the instrument. Figure 1.3 presents the state averages of the periodic fields of Figure 1.1 and the fractal fields of Figure 1.2. A porous medium may possess both fractal characteristics and natural scales.

Heterogeneity may also be observed in different sampling locations and different sampling frequencies. This may be measured explicitly by multiple experiments over a field with a constant window of observation and resolution. Still, if the samples show the same results, heterogeneity may be implicitly present due to scale constants of the experimental equipment or the choice of the sample locations. Only if the correlation scale is much less than the size of the flow domain, statistical stationarity prevails.

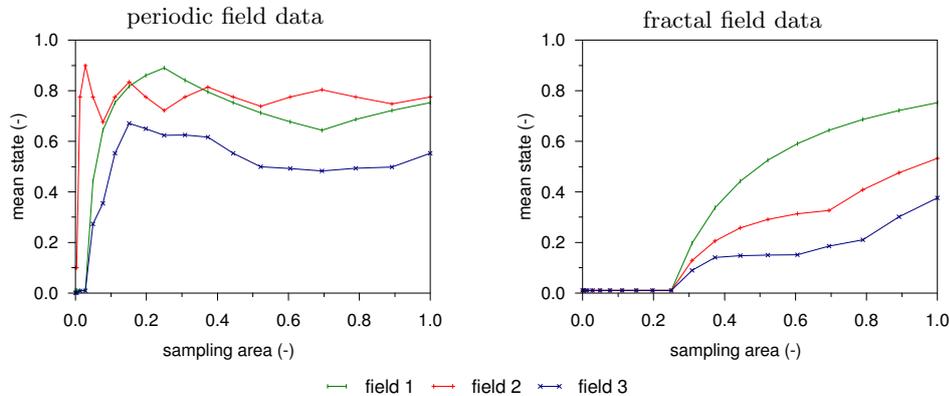


Figure 1.3: State average for periodic and fractal fields.

For these conditions, a statistical framework of porous medium variables can be built by random space functions. Then the actual porous medium corresponds to a single realization of the random space functions ensemble. The space average is often captured by the assemble average under the ergodic hypothesis [33]. As measuring devices that provide a fully resolved deterministic geologic description of the subsurface are not available, the field has to be based on a set of measurements at different locations. Geo-statistical methods then generate the geological model on field scale. Laboratory experiments provide measurements of the subunit parameters at a fine scale. If laboratory measurements are treated stochastically within the geological model, then the structural model of the subsurface should be built on the same scale as these supporting measurements. Geo-statistical methods generate flow problems with strongly discontinuous coefficients. Figure 1.4 gives an example of the intrinsic permeability variation on field scale. At the present, fully resolved flow simulations are computationally intractable at this resolution, and approximated flow problems are solved instead using upscaling techniques or more general multiscale methods. Solving the upscaled flow problems still requires an efficient solution technique like multigrid or adaptive gridding.

Well established commercially available codes for solving partly saturated flow problems coupled to solute transport and heat transport in the subsoil include: Sutra [118], Hst3D [78], Feflow [39] and Plaxflow [20]. All of them simulate both steady state and transient flow in a three-dimensional domain. None of these codes however, are capable of dealing with heterogeneous permeability fields which show a large contrast, and most of them operate on fixed meshes or grids.

### Upscaling techniques

The development of upscaling techniques, particularly in petroleum reservoir engineering, has been motivated by geo-statistical reservoir description algorithms [6, 8, 42, 43, 64, 96, 97, 23]. Nowadays, upscaling methods are applied that aim to propagate information over the hierarchy of scales both functionally and structurally from fine

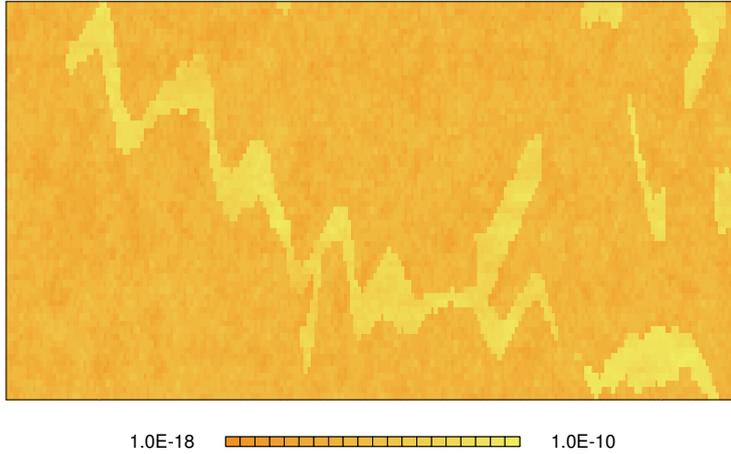


Figure 1.4: Intrinsic permeability field ( $\text{m}^2$ ) for layer 36 of the tenth SPE comparative solution project [26].

scale to coarse scale and not vice versa. In subsurface flow simulations the permeability is the dominant parameter and a highly heterogeneous parameter affecting the subsurface flow at field scale. Permeability is not an additive parameter so it is not possible to calculate an equivalent coarse scale permeability as a simple average of fine scale measurements. A complete equivalence between the heterogeneous medium and the artificial homogeneous medium is impossible. Therefore a flow criterion or a criterion on the energy dissipation often defines an equivalent permeability. Only if the scale of variation is much smaller than the coarse scale then the equivalent permeability matches the effective permeability. The effective permeability is a constant second order tensor variable on the coarse scale, and applies to a medium that is statistically homogeneous on the large scale [101]. Effective permeability only holds for discrete hierarchical systems where scales can be decomposed. On a finite-size block the concept of statistical homogeneity does not hold if the observed block is too small. Here the upscaled permeability or block permeability is the equivalent permeability of this finite-size volume [101]. The upscaled permeability is non-unique and depends on the boundary conditions of the sample domain.

An overview of scaling techniques can be found in literature [30, 101]. Power averaging, renormalization and homogenization will be mentioned here. Power averaging generalizes simple averages like arithmetic mean, geometric mean and harmonic mean. A more advanced averaging technique is renormalization. This technique locally computes a scalar permeability coefficient by treating the medium as a resistance network [76, 77, 54]. The method of homogenization assumes a periodic structure of the permeability coefficients [67, 94]. The effective permeability then follows from a local flow computation on a domain subjected to periodic boundary conditions. Numerical homogenization techniques either apply local scaling [129, 42, 7] or global scaling [16, 65]. In addition global-local scaling techniques [24, 56, 86, 44] first compute local averages

and then replace them by global flow computation averages, if necessary, to obtain a predefined degree of accuracy. Zijl and Trykozko [129] proposed a local numerical volume averaging technique, which calculates equivalent permeability tensors. They pointed out that their pressure-dissipation approach has mathematical advantages, such as the possibility to use it in combination with multigrid methods.

For all these scaling techniques, the permeability coefficient needs to be resolved on the finest grid. If the field is under-resolved then numerical homogenization results depend on the size of the finest grid as the discretization on the finest grid introduces some kind of averaging already. For numerical homogenization techniques over-resolved computations increase the accuracy of the predicted equivalent permeability.

### Multiscale finite element method

Multiscale finite element methods are designed to capture the large scale behavior of the solution without resolving the fine scale features [81, 68, 69, 24, 73, 124, 2, 63, 3, 22, 27, 46, 63, 75, 80, 87, 91, 74]. These methods capture the small scale behavior within each element by multiscale finite element basis functions. The modified basis functions need to be compatible, which means that they need to be  $C^0$ -continuous across the element boundaries and they should satisfy the homogeneous flow equation (no source term) for each coarse scale element. Element wise local boundary conditions mostly follow from a dimensionally reduced problem, as linear boundary conditions often produce less accurate algorithms. The small-scale information is transferred to the large-scale by integration over the small-scale elements and constructing a large-scale element matrix. This way, the method reduces the size of the computation. For a two-dimensional problem the operation count is about twice that of a conventional finite element method. However, the construction of the basis functions is fully decoupled between elements and is naturally adapted to parallel computing. The method is not restricted to assumptions on the media, such as scale separation and periodicity. Moreover, the number of scales involved is irrelevant to the computational cost. In its original form, the two-level method can be seen as a domain decomposition technique [98]. The procedure removes fine scale nodes from the subdomain, but introduces errors at the subdomain boundaries. These errors are not corrected iteratively as in standard domain decomposition methods.

Multiscale finite elements do not calculate effective permeability tensors explicitly, but implicitly capture the fine-scale permeability distribution. MacLachlan [92] related multigrid upscaling of governing equations to multiscale finite element discretization. He obtained multiscale basis function weights recursively by directional lumping and variational coarsening.

### Multigrid solver

Multigrid solvers are able to obtain optimal convergence properties but have to be adjusted for problems with strongly discontinuous coefficient fields [5, 37, 38, 28, 34, 103, 102, 120, 45]. Multigrid methods are mostly used in combination with finite difference techniques, although the variational finite element framework applies to multigrid more naturally [25, 90]. The methods apply a series of coarser grids to accelerate convergence

of basic relaxation methods. For problems with locally highly varying or discontinuous coefficients functions, the performance and robustness of a multigrid solver depends highly on the choice of the coarse grid operators. In the best case, computational work and memory requirements scale linearly with the number of unknowns. There are two classes of multigrid methods: geometric multigrid and algebraic multigrid. Geometric multigrid operates on a hierarchy of grids, whereas algebraic multigrid generates a hierarchy of algebraic equations. Both types apply a discretization of the flow problem on the finest grid. The discretization of the flow equation can be interpreted as an averaging or filtering in itself. In the geometric framework, homogenization methods provide equivalent material behavior on a coarse grid. Flow equations are then discretized on coarser grids to provide coarse grid operators. In the algebraic framework variational coarsening provides a discretization on successive levels. Matrix dependent interpolation supports variational coarsening, and aims to provide coarse grid operators that provide good convergence rates for the multigrid algorithm. As an alternative, Schur complements of the approximated fine-scale matrix with respect to the coarser scale can be used. Variational coarsening implicitly generates multiscale basis functions.

Multigrid upscaling of the permeability field was used by Griebel [59] and Knappek [82]. They applied variational coarsening to calculate coarse-scale equations and extracted an equivalent permeability tensor by interpreting the stencil of the multigrid homogenized operator as a linear combination of the finite difference stencils associated with second order derivatives. As Griebel and Knappek pointed out, matrix dependent prolongation and Schur complement approximations lead to energy dependent averaging procedures and to averaged equations. Moulton [95] was able to obtain the equivalent permeability tensor directly. All authors apply a lumping procedure to calculate the intergrid transfer operators used in the variational coarsening, which was proposed by Dendy [38].

### **Adaptive formulation**

Adaptive grid refinement concentrates the computational work on localized phenomena like sinks and concentration fronts [1, 11, 17, 49, 58, 93, 109, 126, 36, 35]. Figure 1.5 shows the flexibility of the finite element method with this respect. The refinement technique proposed in this thesis divides a coarse line element in two fine line elements, a triangular element splits into four child elements and a tetrahedral element generates eight child elements. Successive refinement of a line element over five levels generates 16 subelements out of a single line element at the base level. A single triangular element subdivides over five levels into 256 subelements and a tetrahedral parent element generates 4096 child elements on the fine level. These numbers show that an adaptive grid refinement technique that concentrates the computational work on localized phenomena reduces the computation effort to a large extent, especially in three dimensions.

### **Scope of the work**

The aim of this work is to provide a fast and robust algorithm to simulate groundwater flow, solute transport and heat transport through highly heterogeneous porous media.

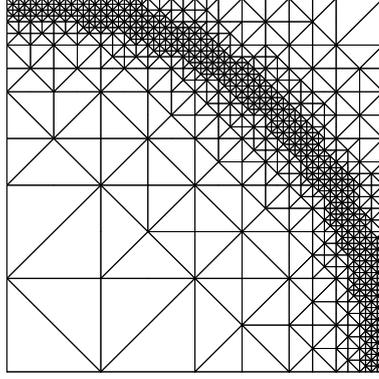


Figure 1.5: A finite element mesh with refinements.

Flow and transport simulations will be carried out on a geo-statistically generated geological model at field scale supported by laboratory scale measurements. The flow domain may either be saturated or (partly) unsaturated. As a result of changes in solute mass fraction and temperature the density and viscosity of the fluid may vary in time. The hydraulic permeability tensor depends on the intrinsic permeability tensor, the relative permeability, the viscosity of the fluid, the density of the fluid and the gravitation constant. The algorithm should be able to capture both the steady state and time-dependent fine scale character of the hydraulic permeability tensor. Though constitutive relations may change for different scales, Darcy's law is supposed to remain valid on both laboratory scale and field scale. The algorithm had to be based on the nodal conforming finite element method because this method is widely used to solve geotechnical problems as it is able to capture complex geometries. The algorithm should support parallel computing and adaptive grid refinement.

### Outline of the thesis

The outline of this thesis follows the modeling process. First the physical behavior of the subsurface system is captured by a mathematical model. This model translates to a numerical model which has to be solved. The behavior of the numerical model will be verified for a number of known partly homogeneous benchmark problems. Finally, the model will be applied to resolve two problems that focus on flow and transport through a highly heterogeneous subsurface.

Chapter 2 presents the mathematical model as a set of coupled partial differential equations. Continuum mechanics at microscopic level produces microscopic balance equations and the method of volume averaging derives the macroscopic balance equations for mass, momentum and energy. Equations of state relate the unknowns in the balance equations to primary unknowns. Both empirical and more fundamental constitutive relations for porosity, degree of saturation, permeability, fluid density and fluid viscosity are presented. The resulting partial differential equations are subjected to boundary conditions and initial conditions for the fully coupled subsurface flow, solute

transport and heat transport problem.

Chapter 3 gives the numerical model as a set of algebraic equations. The conformal finite element method first discretizes the partial differential equations of the mathematical model in space. The finite element method applies to the basic equations for flow, solute transport and heat transport separately. An implicit time integration scheme integrates the ordinary differential equation of the problems. Nonlinearities in the equations are resolved sequentially by Picard iterations. Finally the fully coupled resulting sets of linearized algebraic equations are presented, together with a procedure for generation of initial conditions and handling of linear and geometrically nonlinear boundary conditions.

Chapter 4 presents the newly developed adaptive multiscale finite element method and its derived multigrid solver. The adaptive formulation and the multiscale formulation will be explained separately. Discrete adaptivity formulates mesh refinement, and functional adaptivity facilitates coarsening of the mesh. Multiscale basis functions extrapolate the coarse scale results to the fine scale and construct intergrid transfer operators. The coarse mesh equations apply to the adaptively refined mesh and the coarse grid operators apply to the sequence of coarse grids, which support the multigrid solver.

Chapter 5 covers the model verification part on piecewise homogeneous domains. The first section focuses on saturated and unsaturated flow and considers stable saturated flow through a dam and potentially unstable unsaturated furrow infiltration. The second section focuses on solute transport and heat transport and covers stable density driven solute transport for a salt tongue that intrudes a fresh water domain and unstable convective heat transport in a domain heated from below. Adaptive results are compared with results of standard fine scale conforming finite element calculations.

Chapter 6 includes two applications of the adaptive multiscale finite element method on highly heterogeneous domains. The first section includes a simplified geothermal storage process in a shallow marine system. The second application focuses on geothermal energy production in a deep fluvial system.

Chapter 7 contains the conclusions.

# Chapter 2

## Mathematical model

The subsurface is considered to be a porous medium [12, 116, 51, 33, 9, 52]. A porous medium may consist of a single solid phase and multiple fluid phases. Phases are material subdomains that are separated from each other by phase interfaces. Each phase may consist of a number of miscible chemical species. In the present approach the fluid phases are considered to be immiscible. The mathematical model of this multi-phase multi-species system, is based on fundamental physical principles: mass conservation, momentum conservation, and energy conservation [85, 108]. The balance laws are firstly formulated on a microscopic level by an Eulerian or spatial description. A spatial averaging procedure transforms the microscopic balance equations over representative elementary volumes to the macroscopic level [67, 123]. At the macroscopic scale, the phases are described as overlapping continua, where macroscopic physical laws and conservation principles apply. Corresponding constitutive relations express unknowns in the governing balance equations on the macroscale in measurable parameters. These relations often have an empirical nature. Equations of state give fundamental thermodynamic relationships. This thesis focuses on a three-phase system, which consists of a solid phase (soil skeleton), a liquid phase (groundwater) and a gas phase (air). The model is restricted to single phase flow as the gas-phase is considered to be stagnant and the solid phase is assumed to deform relatively slowly.

Section 2.1 presents the macroscopic balance equations for conservation of mass, conservation of momentum and conservation of energy. Balance equations apply to the solid phase, the liquid phase and the liquid phase components. Section 2.2 gives the constitutive equations that relate unknowns in the balance equation like porosity, saturation, permeability, density and viscosity to a set of independent primary unknowns; pressure, mass fraction and temperature. Section 2.3 states the mathematical problem which needs to be solved subject to properly defined initial conditions and boundary conditions. Characteristic numbers describe the flow regime and support the dimensional analysis. Throughout this section, the coordinate free tensor notation will be used, which is outlined in appendix A. This appendix also introduces the concept of continuum mechanics and the method of volume averaging.

## 2.1 Balance equations

Macroscopic balance equations for conservation of mass, conservation of momentum and conservation of energy form the basis of the mathematical flow and transport model. The momentum conservation equation for the liquid phase supports the constitutive law of Darcy. The momentum conservation equation for a constituent of the liquid phase supports an Fickian type equation for non-convective species flux. The macroscopic energy balance equation for the complete system assumes thermal equilibrium between all phases.

Section 2.1.1 presents the mass balance equations, which prescribe mass conservation of the solid phase, mass conservation of the fluid phase and mass conservation of chemical constituents of the liquid phase. Section 2.1.2 gives the momentum balance equations, which involve momentum conservation of the liquid phase and momentum conservation of the liquid phase chemical components. Energy conservation for the total system gives the energy balance equation presented by section 2.1.3.

### 2.1.1 Mass balance equation

The microscopic equation for conservation of mass is given by equation (A.19) and the macroscopic mass conservation equation for a chemical species in the fluid phase is given by equation (A.35). This equation holds for all chemical components of all phases of the porous medium. Diersch and Kolditz [41] presented the macroscopic mass balance equation for a multi-phase, multi-species system as

$$\frac{\partial}{\partial t} (\phi^\alpha \rho^\alpha \omega_k^\alpha) + \nabla \cdot (\phi^\alpha \rho^\alpha \omega_k^\alpha \mathbf{v}^\alpha) + \nabla \cdot \mathbf{j}_k^\alpha = M_k^\alpha + R_k^\alpha, \quad (2.1)$$

where  $\alpha$  is the phase indicator and  $k$  relates the chemical component or miscible species of the phase. For the porous medium under investigation  $\alpha \in (l, g, s)$  where  $l$  denotes the liquid phase,  $g$  indicates the gas phase, and  $s$  is the solid phase of the porous medium. In the left hand side of equation (2.1)  $t$  [t] denotes time,  $\phi^\alpha$  [-] is the volume fraction of the  $\alpha$ -phase,  $\rho^\alpha$  [ml<sup>-3</sup>] is the phase density,  $\omega_k^\alpha$  [-] is the mass fraction of species  $k$  in the  $\alpha$ -phase,  $\mathbf{j}$  [ml<sup>-2</sup>t<sup>-1</sup>] is the non-convective species flux vector due to dispersion or diffusion, and  $\mathbf{v}$  [lt<sup>-1</sup>] is the phase velocity vector. In the right hand side  $R^\alpha$  [ml<sup>-3</sup>t<sup>-1</sup>] is the chemical reaction rate term and  $M^\alpha$  [ml<sup>-3</sup>t<sup>-1</sup>] denotes the bulk solute mass source term for internal and external transfer. The mass source term will have a negative value if mass is extracted from the porous medium. The sum of the mass fraction over all species in the phase equals one, and the sum of the non-convective terms and chemical rate terms equals zero, according to

$$\sum_{\alpha}^{n^\alpha} \phi^\alpha = 1, \quad \sum_k^{n_k^\alpha} \omega_k^\alpha = 1, \quad \sum_{\alpha}^{n^\alpha} \mathbf{j}_k^\alpha = 0, \quad \sum_{\alpha}^{n^\alpha} R_k^\alpha = 0. \quad (2.2)$$

Using these expressions, the mass balance equation for each phase can be written as

$$\frac{\partial}{\partial t} (\phi^\alpha \rho^\alpha) + \nabla \cdot (\phi^\alpha \rho^\alpha \mathbf{v}^\alpha) = M^\alpha. \quad (2.3)$$

By definition the total mass source term reads  $\sum_k M_k^\alpha = M^\alpha$ .

### Solid phase mass balance equation

The mass balance for the solid phase follows from the macroscopic mass balance equation (2.3) and the definition of the solid phase volume fraction given by equation (A.30) as

$$\frac{\partial}{\partial t} ((1-n)\rho^s) + \nabla \cdot ((1-n)\rho^s \mathbf{v}^s) = 0, \quad (2.4)$$

where the source term for the solid phase has not been included, so  $M^s = 0$ . For incompressible solid grains, the density of the solid phase  $\rho^s$  is constant and the equation simplifies to

$$\frac{\partial n}{\partial t} - (1-n)\nabla \cdot \mathbf{v}^s + \mathbf{v}^s \cdot \nabla n = 0. \quad (2.5)$$

An indicative value for the solid phase density of sand is  $2.65 \cdot 10^3 \text{ kg/m}^3$ . In general, porous media deform slowly and  $\mathbf{v}^s \cdot \nabla n \approx 0$ , so the mass balance equation for the solid phase becomes

$$\frac{\partial n}{\partial t} = (1-n)\nabla \cdot \mathbf{v}^s. \quad (2.6)$$

Here  $\nabla \cdot \mathbf{v}^s$  is the volume strain rate of the solid matrix.

### Liquid phase mass balance equation

The liquid phase mass balance equation follows from equation (2.3) and the definition of the liquid phase volume fraction given by equation (A.31), and reads

$$\frac{\partial}{\partial t} (nS\rho^l) + \nabla \cdot (nS\rho^l \mathbf{v}^l) = M^l. \quad (2.7)$$

Saturation reads  $S = V^l/V^p$ , where  $V^l$  is the volume of the liquid phase and  $V^p$  is the total pore volume. According to the definition of the filtration velocity given by equation (A.37) and the definition of the liquid phase volume fraction presented by equation (A.31), the macroscopic velocity vector of the liquid phase reads

$$\mathbf{v}^l = \frac{\mathbf{q}^l}{nS} + \mathbf{v}^s, \quad (2.8)$$

where  $\mathbf{q}^l$  is the relative (Darcy) bulk velocity of the liquid phase [ $\text{lt}^{-1}$ ]. Rewriting the mass balance equation in terms of the filtration velocity and expanding the time derivative term yields

$$S\rho^l \frac{\partial n}{\partial t} + n\rho^l \frac{\partial S}{\partial t} + nS \frac{\partial \rho^l}{\partial t} + \nabla \cdot (\rho^l \mathbf{q}^l) + nS\rho^l \nabla \cdot \mathbf{v}^s + \mathbf{v}^s \cdot \nabla (nS\rho^l) = M^l. \quad (2.9)$$

Using the definition for the divergence of the solid phase macroscopic velocity given by equation (2.6), the liquid phase mass balance equation reads

$$\frac{S\rho^l}{1-n} \frac{\partial n}{\partial t} + n\rho^l \frac{\partial S}{\partial t} + nS \frac{\partial \rho^l}{\partial t} + \nabla \cdot (\rho^l \mathbf{q}^l) = M^l, \quad (2.10)$$

where for slowly deformable media and slightly compressible fluids it was assumed that  $\mathbf{v}^s \cdot \nabla (nS\rho^l) \approx 0$ .

### Liquid phase component mass balance equation

The macroscopic mass balance equation for a single species  $k$  in the liquid phase follows from equation (2.1) and the definition of the liquid phase volume fraction given by equation (A.31). The liquid phase component mass balance equation then reads

$$\frac{\partial}{\partial t} (nS\rho^l\omega_k^l) + \nabla \cdot (nS\rho^l\omega_k^l\mathbf{v}^l) + \nabla \cdot \mathbf{j}_k^l = M_k^l + R_k^l. \quad (2.11)$$

Expanding the time derivative term and inserting the expression for the macroscopic liquid phase velocity presented by equation (2.8) gives

$$\begin{aligned} S\rho^l\omega_k^l \frac{\partial n}{\partial t} + n\rho^l\omega_k^l \frac{\partial S}{\partial t} + nS\rho^l \frac{\partial \omega_k^l}{\partial t} + nS\omega_k^l \frac{\partial \rho^l}{\partial t} + \nabla \cdot (\rho^l\omega_k^l\mathbf{q}^l) \\ + nS\rho^l\omega_k^l \nabla \cdot \mathbf{v}^s + \mathbf{v}^s \cdot \nabla (nS\rho^l\omega_k^l) + \nabla \cdot \mathbf{j}_k^l = M_k^l + R_k^l. \end{aligned} \quad (2.12)$$

For slowly deformable media and slightly compressible fluids  $\mathbf{v}^s \cdot \nabla (nS\rho^l\omega_k^l) \approx 0$ . Inserting equation (2.6) which gives an expression for the macroscopic velocity, the solute mass balance equation yields

$$\begin{aligned} \frac{S\rho^l\omega_k^l}{1-n} \frac{\partial n}{\partial t} + n\rho^l\omega_k^l \frac{\partial S}{\partial t} + nS\rho^l \frac{\partial \omega_k^l}{\partial t} + nS\omega_k^l \frac{\partial \rho^l}{\partial t} + \nabla \cdot (\rho^l\omega_k^l\mathbf{q}^l) \\ + \nabla \cdot \mathbf{j}_k^l = M_k^l + R_k^l. \end{aligned} \quad (2.13)$$

This form is known as the divergent form of the transport equation. The convective form follows from extracting  $\omega_k^l$  times the liquid phase mass balance equation (2.10) from the divergent transport equation (2.13). The convective form reads

$$nS\rho^l \frac{\partial \omega_k^l}{\partial t} + \rho^l\mathbf{q}^l \cdot \nabla \omega_k^l + \nabla \cdot \mathbf{j}_k^l = M_k^l - M^l\omega_k^l + R_k^l, \quad (2.14)$$

where  $M^l$  [ $\text{ml}^{-3}\text{t}^{-1}$ ] denotes the liquid source term, and  $M_k^l$  [ $\text{ml}^{-3}\text{t}^{-1}$ ] expresses the solute mass entering the porous medium.

#### 2.1.2 Momentum balance equation

The microscopic equation for conservation of momentum is given by equation (A.20). Diersch and Kolditz [41] presented the general macroscopic momentum balance equation for a phase as

$$\frac{\partial}{\partial t} (\phi^\alpha \rho^\alpha \mathbf{v}^\alpha) + \nabla \cdot (\phi^\alpha \rho^\alpha \mathbf{v}^\alpha \mathbf{v}^\alpha) - \nabla \cdot \boldsymbol{\sigma}^\alpha = \phi^\alpha \rho^\alpha \mathbf{g} + \phi^\alpha \mathbf{f}^\alpha, \quad (2.15)$$

where  $\phi^\alpha \rho^\alpha \mathbf{g}$  [ $\text{lt}^{-2}$ ] represents an external body force vector due to gravity only as ionic attractions are neglected,  $\mathbf{f}^\alpha$  [ $\text{ml}^{-2}\text{t}^{-2}$ ] denotes an internal drag vector or momentum exchange vector, and  $\boldsymbol{\sigma}^\alpha$  [ $\text{ml}^{-1}\text{t}^{-2}$ ] denotes the viscous stress tensor if a fluid is considered or the Cauchy stress tensor for the case of a solid phase.

For a fluid in motion, the stress tensor is decomposed in two parts and is written as

$$\boldsymbol{\sigma}^\alpha = -\phi^\alpha p^\alpha \mathbf{I} + \phi^\alpha (\boldsymbol{\sigma}^\alpha)^D. \quad (2.16)$$

The deviatoric part of the stress tensor  $(\boldsymbol{\sigma}^\alpha)^D$  depends on the rate of deformation, expressed by  $\nabla \mathbf{v}^\alpha$ , the velocity gradient and the dynamic viscosity of the fluid  $\mu^\alpha$  [ $\text{ml}^{-1}\text{t}^{-1}$ ]. The pressure  $p$  [ $\text{ml}^{-1}\text{t}^{-2}$ ] does not depend on the rate of deformation. A constitutive equation expresses the inter-facial drag term of momentum exchange for a fluid phase in measurable flow quantities. For linear drag the Darcy term is imposed as

$$\mathbf{f}^\alpha = -\mu^\alpha (\mathbf{K}^\alpha)^{-1} \cdot \mathbf{q}^\alpha, \quad (2.17)$$

where  $(\mathbf{K}^\alpha)^{-1}$  denotes the inverse of the intrinsic permeability tensor  $\mathbf{K}$  [ $l^2$ ] and  $\mathbf{q}^\alpha$  [ $l\text{t}^{-1}$ ] is the volumetric flux density. This constitutive relation can be extended with a quadratic drag term (to model turbulent flow), known as the Forchheimer term, to hold

$$\mathbf{f}^\alpha = -\mu^\alpha (\mathbf{K}^\alpha)^{-1} \cdot \mathbf{q}^\alpha - C_f q^\alpha \mu^\alpha (\mathbf{K}^\alpha)^{-1} \cdot \mathbf{q}^\alpha. \quad (2.18)$$

Here  $C_f$  [ $l^{-1}\text{t}$ ] is the Forchheimer coefficient, and  $q^\alpha$  [ $l\text{t}^{-1}$ ] is the magnitude of the volumetric flux density  $\mathbf{q}^\alpha$ .

Leijnse [88] extracted an expression for the non-convective flux from the momentum equation of a constituent in the fluid phase. In this section the resulting phenomenological law for the non-convective species fluxes  $\mathbf{j}_k^\alpha$  will be given. This Fickian type equation reads

$$\mathbf{j}_k^\alpha = -\rho^\alpha \mathbf{D}^\alpha \cdot \nabla \omega_k^\alpha, \quad (2.19)$$

where  $\mathbf{D}^\alpha$  [ $l^2\text{t}^{-1}$ ] is the second rank hydrodynamic dispersion tensor. According to De Josselin de Jong [32] and Scheidegger [105] hydrodynamic dispersion consists of molecular diffusion and mechanical dispersion. For an isotropic porous medium the hydrodynamic dispersion tensor reads

$$\mathbf{D}^\alpha = \phi^\alpha D_c^\alpha \mathbf{T} + \phi^\alpha \mathbf{D}_m^\alpha, \quad \mathbf{D}_m^\alpha = \alpha_t v^\alpha \mathbf{I} + (\alpha_l - \alpha_t) \frac{\mathbf{v}^\alpha \mathbf{v}^\alpha}{v^\alpha}, \quad (2.20)$$

where  $v^\alpha = \|\mathbf{v}^\alpha\|_2$  [ $l\text{t}^{-1}$ ] is the magnitude of the phase flux vector,  $\mathbf{v}\mathbf{v}$  is the dyadic-product of the velocity vector  $\mathbf{v}$ ,  $D_c$  [ $l^2\text{t}^{-1}$ ] denotes the coefficient of molecular diffusion, and  $\mathbf{T}$  [ $-$ ] is the tortuosity tensor. The mechanical dispersion is written as  $\mathbf{D}_m$  [ $l^2\text{t}^{-1}$ ],  $\mathbf{I}$  [ $-$ ] denotes the unit tensor,  $\alpha_l$  [ $l$ ] is the longitudinal dispersivity and  $\alpha_t$  [ $l$ ] is the transverse dispersivity. Both transverse and longitudinal dispersivity are length scale dependent due to the multiscale nature of the subsurface. Dispersivity parameters apply to an equivalent homogeneous porous medium.

Leijnse, Hassanizadeh and Schotting [88, 61, 62, 106] proposed a nonlinear extension that captures large concentration gradients. This implicit formulation of the non-convective species mass flux reads

$$\mathbf{j}_k^\alpha + C_H j_k^\alpha \mathbf{j}_k^\alpha = -\rho^\alpha \mathbf{D}^\alpha \cdot \nabla \omega_k^\alpha, \quad (2.21)$$

where  $j_k^\alpha = \|\mathbf{j}_k^\alpha\|_2$  [ $\text{ml}^{-2}\text{t}^{-1}$ ] is the magnitude of the species mass flux vector and  $C_H$  [ $\text{m}^{-1}l^2\text{t}$ ] is known as the non-Fickian high-concentration dispersion coefficient.

### Liquid phase momentum balance equation

The liquid phase momentum balance equation results from equation (2.15) and the expression for the volume fraction of the liquid phase given by equation (A.31). Conservation of momentum for the liquid phase is expressed by

$$\frac{\partial}{\partial t} (nS\rho^l \mathbf{v}^l) + \nabla \cdot (nS\rho^l \mathbf{v}^l \mathbf{v}^l) - \nabla \cdot \boldsymbol{\sigma}^l = nS\rho^l \mathbf{g} + nS\mathbf{f}^l. \quad (2.22)$$

Disregarding the inertial terms  $\frac{\partial}{\partial t} (nS\rho^l \mathbf{v}^l)$  for slow movement of the fluid, excluding the compression and deformation work, by  $\nabla \cdot (\rho^l \mathbf{v}^l \mathbf{v}^l) \approx 0$ , inserting the expression for the viscous stress tensor given by equation (2.16) and the internal drag term expressed by equation (2.18) gives

$$\begin{aligned} nS\nabla p^l + p^l \nabla (nS) - nS\nabla \cdot (\boldsymbol{\sigma}^l)^D - (\boldsymbol{\sigma}^l)^D \nabla (nS) \\ = nS\rho^l \mathbf{g} - nS\mu^l (\mathbf{K}^l)^{-1} \cdot \mathbf{q}^l - C_f nS q^l \mu^l (\mathbf{K}^l)^{-1} \cdot \mathbf{q}^l. \end{aligned} \quad (2.23)$$

Replacing  $\nabla \cdot (\boldsymbol{\sigma}^l)^D$  by the Brinkman term  $\mu^l \nabla^2 \mathbf{q}^l$  where  $\nabla^2 = \nabla \cdot \nabla$ , the momentum equation for the liquid phase reads

$$\nabla p^l - \mu^l \nabla^2 \mathbf{q}^l = \rho^l \mathbf{g} - \mu^l (\mathbf{K}^l)^{-1} \cdot \mathbf{q}^l - C_f q^l \mu^l (\mathbf{K}^l)^{-1} \cdot \mathbf{q}^l, \quad (2.24)$$

or

$$\mathbf{q}^l + C_f q^l \mathbf{q}^l = -\frac{\mathbf{K}^l}{\mu^l} \cdot (\nabla p^l - \rho^l \mathbf{g}) + \mathbf{K}^l \cdot \nabla^2 \mathbf{q}^l. \quad (2.25)$$

This equation was derived by Leijnse [88] for saturated flow conditions. For unsaturated flow the term  $\nabla (nS)$  is small. Without the Forchheimer and Brinkman effects the fluid momentum equation is then given by

$$\mathbf{q}^l = -\frac{\mathbf{K}^l}{\mu^l} \cdot (\nabla p^l - \rho^l \mathbf{g}). \quad (2.26)$$

Here  $\mathbf{q}^l$  [lt<sup>-2</sup>] is the volumetric flux density,  $\mathbf{K}^l$  [l<sup>2</sup>] is the second rank (intrinsic) permeability tensor of the porous medium,  $\mu$  [ml<sup>-1</sup>t<sup>-1</sup>] denotes the dynamic viscosity of the liquid,  $p$  [ml<sup>-1</sup>t<sup>-2</sup>] denotes the pressure, and  $\mathbf{g}$  [lt<sup>-2</sup>] is the gravitational body force vector. Equation (2.26) is known as Darcy's law. This empirical equation will be used for saturated soils. A phenomenologic law will extend Darcy's law for unsaturated flow conditions.

### Liquid phase component momentum balance equation

The Fickian type equation (2.19) for the non-convective species flux in the liquid phase reads

$$\mathbf{j}_k^l = -\rho^l \mathbf{D}^l \cdot \nabla \omega_k^l. \quad (2.27)$$

The hydrodynamic dispersion tensor for the liquid phase reads

$$\mathbf{D}^l = nSD_c^\alpha \mathbf{T} + \mathbf{D}_m^l, \quad \mathbf{D}_m^l = \alpha_t q^l \mathbf{I} + (\alpha_l - \alpha_t) \frac{\mathbf{q}^l \mathbf{q}^l}{q^l}. \quad (2.28)$$

Here equation (2.8) is approximated by  $\mathbf{v}^l \approx \mathbf{q}^l/nS$ , excluding solid matrix velocity effects.

### 2.1.3 Energy balance equation

The local microscopic equation for conservation of energy was given by equation (A.21). Diersch and Kolditz [41] presented the phase related macroscopic energy balance equation as

$$\begin{aligned} \frac{\partial}{\partial t} \left[ \phi^\alpha \rho^\alpha \left( E_h^\alpha + \frac{v^\alpha v^\alpha}{2} \right) \right] + \nabla \cdot \left[ \phi^\alpha \rho^\alpha \mathbf{v}^\alpha \left( E_h^\alpha + \frac{v^\alpha v^\alpha}{2} \right) \right] \\ + \nabla \cdot \mathbf{j}_h^\alpha - \nabla \cdot (\boldsymbol{\sigma}^\alpha \cdot \mathbf{v}^\alpha) = \phi^\alpha \rho^\alpha \mathbf{g} \cdot \mathbf{v}^\alpha + \phi^\alpha \mathbf{f}^\alpha \cdot \mathbf{v}^\alpha + H^\alpha, \end{aligned} \quad (2.29)$$

where  $\mathbf{j}_h$  [mt<sup>-3</sup>] expresses the non-convective (heat) energy flux,  $E_h$  [l<sup>2</sup>t<sup>-2</sup>] denotes the internal (thermal) energy,  $v^\alpha$  [lt<sup>-1</sup>] is the magnitude of the macroscopic phase velocity vector  $\mathbf{v}^\alpha$ , and  $H^\alpha$  [ml<sup>-1</sup>t<sup>-3</sup>] accounts for the energy source, which captures internal and external heat supply. For subsurface flow kinetic energy effects are usually disregarded, so  $\partial(v^\alpha v^\alpha)/\partial t \approx 0$ ,  $\nabla \cdot (v^\alpha v^\alpha) \approx 0$ ,  $\nabla \cdot (\boldsymbol{\sigma}^\alpha \cdot \mathbf{v}^\alpha) \approx 0$ ,  $\phi^\alpha \rho^\alpha \mathbf{g} \cdot \mathbf{v}^\alpha \approx 0$ , and  $\phi^\alpha \mathbf{f}^\alpha \cdot \mathbf{v}^\alpha \approx 0$ . These assumptions reduce the macroscopic energy balance equation to

$$\frac{\partial}{\partial t} (\phi^\alpha \rho^\alpha E_h^\alpha) + \nabla \cdot (\phi^\alpha \rho^\alpha \mathbf{v}^\alpha E_h^\alpha) + \nabla \cdot \mathbf{j}_\phi^h = H^\alpha. \quad (2.30)$$

The thermodynamic state function for internal energy can be formulated as

$$dE_h^\alpha = \left( T^\alpha \frac{\partial p^\alpha}{\partial T^\alpha} \Big|_{\nu^\alpha, \omega_k^\alpha} - p^\alpha \right) d\nu^\alpha + c^\alpha dT^\alpha + \sum_{k=1}^{n_k^\alpha} \mu_k^\alpha d\omega_k^\alpha, \quad (2.31)$$

where  $\nu^\alpha = 1/\rho^\alpha$  [m<sup>-1</sup>l<sup>3</sup>] is the specific volume,  $\mu_k^\alpha$  [l<sup>2</sup>t<sup>-2</sup>] denotes the chemical potential of species  $k$  and  $c^\alpha$  [l<sup>2</sup>t<sup>-2</sup>T<sup>-1</sup>] is the specific heat capacity. The chemical potential depends on pressure, mass fraction of the constituents and temperature as  $\mu_k^\alpha = \mu_k^\alpha(p^\alpha, \omega_k^\alpha, T^\alpha)$ . The specific heat capacity can be written as

$$c^\alpha = \frac{\partial E_h^\alpha}{\partial T^\alpha} \Big|_{\nu^\alpha, \omega_k^\alpha}. \quad (2.32)$$

If pressure and chemical effects are negligible then the internal energy only depends on temperature. The equation of state for this case can be written as

$$dE_h^\alpha = c^\alpha dT^\alpha. \quad (2.33)$$

For a constant heat capacity, equation (2.33) is reformulated as

$$E_h^\alpha = c^\alpha (T - T_0^\alpha). \quad (2.34)$$

A phenomenological law for the non-convective heat energy flux [48] generalizes the corresponding linear Fourier law as

$$\mathbf{j}_h^\alpha = -\mathbf{H}^\alpha \cdot \nabla T^\alpha, \quad (2.35)$$

where  $\mathbf{H}^\alpha$  [mlt<sup>-3</sup>T<sup>-1</sup>] is the hydrodynamic thermo-dispersivity tensor and  $T$  [T] denotes temperature. Like the hydrodynamic dispersivity tensor, the hydrodynamic thermo-dispersivity tensor consists of a conductive part and a mechanical part. For the solid phase the mechanical part vanishes, and the hydrodynamic dispersivity tensor is expressed as

$$\mathbf{H}^l = nS\lambda^l \mathbf{I} + nSc^l \rho^l \mathbf{D}_m^l, \quad \mathbf{H}^s = (1-n)\lambda^s \mathbf{I}, \quad (2.36)$$

where  $\lambda^l$  and  $\lambda^s$  [mlt<sup>-3</sup>T<sup>-1</sup>] denote heat conductivity of the fluid phase and the solid phase respectively.

### Combined liquid and solid phase energy balance equation

For local thermal equilibrium conditions  $T = T^\alpha$  over all phases, a single energy balance equation replaces the energy balance equations for each phase. Usually the thermal capacity and conductivity of the gas phase is small with respect to the solid and liquid phase so a combined liquid and solid phase energy balance equation captures the conservation of energy in the porous medium. The first term of equation (2.30) and the definition of the internal energy given by equation (2.34) gives the time dependent term in the solid-fluid system as

$$\frac{\partial}{\partial t} (nS\rho^l E_h^l) + \frac{\partial}{\partial t} ((1-n)\rho^s E_h^s) = \frac{\partial}{\partial t} (nSc^l \rho^l T) + \frac{\partial}{\partial t} ((1-n)c^s \rho^s T). \quad (2.37)$$

The second term of the energy balance equation (2.30) includes the liquid phase velocity given by equation (2.8), and specifies the divergence form in the solid-fluid system as

$$\begin{aligned} \nabla \cdot (nS\rho^l \mathbf{v}^l E_h^l) + \nabla \cdot ((1-n)\rho^s \mathbf{v}^s E_h^s) &= \nabla \cdot (c^l \rho^l T \mathbf{q}^l) + nSTc^l \rho^l \nabla \cdot \mathbf{v}^s \\ &+ (1-n)Tc^s \rho^s \nabla \cdot \mathbf{v}^s + \mathbf{v}^s \cdot \nabla (nSTc^l \rho^l) + \mathbf{v}^s \cdot \nabla ((1-n)Tc^s \rho^s \mathbf{v}^s). \end{aligned} \quad (2.38)$$

For slowly deforming media  $\mathbf{v}^s \cdot \nabla (c^l nST\rho^l) \approx 0$  and  $\mathbf{v}^s \cdot \nabla ((1-n)c^s T\rho^s \mathbf{v}^s) \approx 0$ . Using these restriction and substituting the expression for the divergence of the macroscopic solid phase velocity given by equation (2.6), the second term of the combined energy balance equation reads

$$\begin{aligned} \nabla \cdot (nS\rho^l \mathbf{v}^l E_h^l) + \nabla \cdot ((1-n)\rho^s \mathbf{v}^s E_h^s) \\ = \nabla \cdot (\rho^l c^l T \mathbf{q}^l) + \frac{nSTc^l \rho^l}{1-n} \nabla \rho^s \frac{\partial n}{\partial t} + c^s \rho^s T \frac{\partial n}{\partial t}. \end{aligned} \quad (2.39)$$

Assuming the specific heat capacity  $c^s$  and  $c^l$  and the solid phase density  $\rho^s$  to be constant, the combined macroscopic liquid phase and solid phase energy balance equation can be written as

$$\begin{aligned} \frac{c^l ST\rho^l}{1-n} \frac{\partial n}{\partial t} + c^l nT\rho^l \frac{\partial S}{\partial t} + c^l nST \frac{\partial \rho^l}{\partial t} + (c^l nS\rho^l + (1-n)c^s \rho^s) \frac{\partial T}{\partial t} \\ + \nabla \cdot (\rho^l c^l \mathbf{q}^l T) + \nabla \cdot \mathbf{j}_e^l + \nabla \cdot \mathbf{j}_e^s = H^l. \end{aligned} \quad (2.40)$$

This is the divergent form of the combined energy balance equation, also known as the heat transport equation. External heat supply to the liquid phase was introduced by  $H^l$ . Subtracting  $c^l \rho^l$  times the mass balance equation of the liquid phase, given by equation (2.10), produces the convective form of the energy conservation equation for the porous system. The convective form of the energy conservation equation reads

$$(nS\rho^l c^l + (1-n)\rho^s c^s) \frac{\partial T}{\partial t} + (\rho^l c^l \mathbf{q}^l) \cdot \nabla T + \nabla \cdot \mathbf{j}_e^l + \nabla \cdot \mathbf{j}_e^s = H^l - M^l c^l T, \quad (2.41)$$

where the non-convective heat energy fluxes are given by

$$\mathbf{j}_e^l = -\mathbf{H}^l \cdot \nabla T, \quad \mathbf{j}_e^s = -\mathbf{H}^s \cdot \nabla T. \quad (2.42)$$

These two fluxes combine to an average flux, which reads  $\mathbf{j}_e = \mathbf{j}_e^l + \mathbf{j}_e^s$ , as the flux between the two constituents is neglected. This averaged flux reads  $\mathbf{j}_e = -\mathbf{H} \cdot \nabla T$ . The hydrodynamic thermo-dispersion tensors for the liquid and the solid phase are expressed as

$$\mathbf{H} = \mathbf{H}^l + \mathbf{H}^s, \quad \mathbf{H}^l = nS\lambda^l \mathbf{I} + c^l \rho^l \mathbf{D}_m^l, \quad \mathbf{H}^s = (1-n)\lambda^s \mathbf{I}. \quad (2.43)$$

Equation (2.41) expressed energy conservation for the porous medium and equation (2.42) formulated the non-convective heat flux. An indicative value for the specific heat capacity of sandstone is about  $8.4 \cdot 10^2$  J/kgK at 293 K and the thermal conductivity is about 3.5 J/msK at 293 K. The specific heat capacity of pure water is approximately  $4.2 \cdot 10^3$  J/kgK at 293 K and its thermal conductivity is about 0.6 J/msK at 293 K.

## 2.2 Constitutive relations

A number of previously introduced solid material parameters like solid phase density, specific heat capacity, and thermal conductivity are assumed to be constant in time but may vary in space. The specific heat capacity and the thermal conductivity of the liquid phase are considered to be constants in time and space. Longitudinal and transverse dispersivity are assumed to be constant in time but have to capture the hierarchical structure of the heterogeneous subsurface. In most practical applications their values depend on the problem size. In this thesis their value relates to the fine element size as heterogeneous fields are considered explicitly. Molecular diffusion is often neglected as its value is small relative to mechanical dispersion. Constitutive relations for the remaining variables formulate material behavior in terms of measurable quantities. Here fluid pressure, solute mass fraction, and porous medium temperature are chosen as measurable independent variables.

Section 2.2.1 gives an expression for the porosity as a function of fluid pressure. The section presents the equilibrium equation for the solid matrix which results from the momentum equation for the solid matrix. If this equation is decoupled from the flow equation, then it expresses porosity for the case of dissipating water pressures. Empirical relations give a functional dependency for the degree of saturation on the capillary pore pressure and give a relation between relative permeability and saturation.

Section 2.2.2 and section 2.2.3 present the empirical Van Genuchten relations that express saturation and relative permeability as functions of pore pressure. Fluid density and fluid viscosity relate to pressure, solute mass fraction and temperature and their empirical relations will be discussed in section 2.2.4 and section 2.2.5.

### 2.2.1 Porosity

Biot consolidation theory couples the equilibrium equation for the solid skeleton to the storage equation of the fluid. The equilibrium equation follows from the momentum equation of the solid phase, neglects the acceleration terms, and adopts kinematic and constitutive relations. For the solid, kinematic relations express the deformation in terms of strains. Constitutive relations for the material behavior relate stresses to strains. The storage equation follows from the mass balance equation of the fluid and Darcy's law. This subsection proposes a decoupling of both these equations, which gives a relation between porosity and fluid pressure. Concentration and temperature effects are assumed to be negligible and compression work of the skeleton only is taken into account. An elastic material is considered for which the small deformation theory or linear deformation theory applies, as displacements and displacement gradients are small compared to unity.

According to the mass balance equation for the solid phase given by equation (2.6) the rate of change of porosity is equal to the divergence of the macroscopic solid phase velocity. Only the volume of the pores changes as the volume of the solids is considered constant. The relation between solid phase velocity and solid phase displacements can be written as

$$\mathbf{v}^s = \frac{\partial}{\partial t} \mathbf{u}^s, \quad \nabla \cdot \mathbf{v}^s = \frac{\partial}{\partial t} (\nabla \cdot \mathbf{u}^s) = \frac{\partial \varepsilon_v}{\partial t}. \quad (2.44)$$

Substituting this expression into equation (2.6) gives

$$\frac{\partial n}{\partial t} = (1 - n) \frac{\partial \varepsilon_v}{\partial t}, \quad (2.45)$$

where volume strain  $\varepsilon_v$  [-] is the first invariant of the strain tensor  $\boldsymbol{\varepsilon}$  [-] and reads

$$\varepsilon_v = \text{tr}(\boldsymbol{\varepsilon}) = \nabla \cdot \mathbf{u}^s. \quad (2.46)$$

Strains are considered to be positive for extension. The volume strain depends on the stresses in the porous medium.

The macroscopic momentum balance equation (2.15) without acceleration terms resembles the macroscopic static equilibrium of a continuum [116, 19], written as

$$\nabla \cdot \boldsymbol{\sigma} + \boldsymbol{\gamma} = 0, \quad (2.47)$$

where  $\boldsymbol{\sigma}$  [ $\text{ml}^{-1}\text{t}^{-2}$ ] represents the Cauchy (total) stress tensor and  $\boldsymbol{\gamma}$  [ $\text{ml}^{-2}\text{t}^{-2}$ ] denotes the body force vector. Stresses are positive for tensile stress and negative for compression. For a partly saturated soil the weight of the soil is given by  $\boldsymbol{\gamma} = nS\rho^l\mathbf{g} + (1 - n)\rho^s\mathbf{g}$ . The balance of moments requires the Cauchy stress tensor to be symmetrical  $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$ . External forces on the boundary are written as  $\boldsymbol{\tau} = \boldsymbol{\sigma} \cdot \mathbf{n}$ ,

where  $\boldsymbol{\tau}$  [mt<sup>-2</sup>] is the boundary traction vector and  $\mathbf{n}$  [1] is the normal vector pointing out of the domain. Terzaghi's principle states that the total stress sums the liquid pressure in the pores and the effective stress in the soil skeleton. Terzaghi's principle is formulated as

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' - p^l \mathbf{I}, \quad (2.48)$$

where  $\boldsymbol{\sigma}'$  [ml<sup>-1</sup>t<sup>-2</sup>] denotes the effective stress and  $\mathbf{I}$  [-] is the identity tensor. A minus sign applies on the pressure as it is a compressive stress and both effective and total stress are positive for tensile stress. The (effective) stress-strain relationship of an isotropic linear elastic material can be written as

$$\boldsymbol{\sigma}' = \lambda_l \text{tr}(\boldsymbol{\varepsilon}) \mathbf{I} + 2\mu_l \boldsymbol{\varepsilon}, \quad (2.49)$$

where  $\lambda_l$  [ml<sup>-1</sup>t<sup>-2</sup>] and  $\mu_l$  [ml<sup>-1</sup>t<sup>-2</sup>] are called the Lamé constants.

Substitution of the constitutive relation given by equation (2.49) into equation (2.47), which states the static equilibrium of a continuum, gives

$$\gamma + 2\mu_l \nabla \boldsymbol{\varepsilon} + \lambda_l \nabla \text{tr}(\boldsymbol{\varepsilon}) - \nabla p = 0. \quad (2.50)$$

For small deformations the kinematic relation expresses the linear strain tensor in terms of displacements. This relation is written as

$$\boldsymbol{\varepsilon} = \frac{1}{2} \left( (\nabla \mathbf{u}^s)^T + (\nabla \mathbf{u}^s) \right), \quad (2.51)$$

where  $(\nabla \mathbf{u}^s)^T$  is the conjugate of dyad  $(\nabla \mathbf{u}^s)$  and  $\boldsymbol{\varepsilon}$  represents the symmetric part of the dyad  $(\nabla \mathbf{u}^s)$ . The equilibrium equation for the solid matrix follows from equation (2.50) and equation (2.51) as

$$\gamma + (\mu_l + \lambda_l) \nabla \nabla \cdot \mathbf{u}^s + \mu_l \nabla^2 \mathbf{u}^s - \nabla p = 0. \quad (2.52)$$

The expression for the volumetric strain, given by equation (2.46), reformulates equation (2.52) as

$$\gamma + (\mu_l + \lambda_l) \nabla \boldsymbol{\varepsilon}_v + \mu_l \nabla^2 \mathbf{u}^s - \nabla p = 0. \quad (2.53)$$

This equation already relates the strain tensor to the pore water pressure.

For an isotropic material the volume strain is often assumed to be a function of isotropic effective stress. Verruijt [116] stated that in many cases consolidation takes place while the loading of the soil remains constant and it may well be assumed that changes in the total stress will be small. If a soil layer of large extent is loaded by a constant surface load, then the horizontal deformations are negligible and the vertical total stress remains almost constant. For a linear elastic material the relation between vertical effective stress and vertical strain then reads

$$\sigma'_{zz} = (\lambda_l + 2\mu_l) \varepsilon_{zz}. \quad (2.54)$$

The total vertical stress is decomposed into the effective stress and the pore pressure according to  $\sigma_{zz} = \sigma'_{zz} - p$ . As the total stress is assumed to be constant in time, the variation of volumetric strain in time is expressed as

$$\frac{\partial \varepsilon_v}{\partial t} = \alpha_p \frac{\partial p}{\partial t}, \quad (2.55)$$

where  $\alpha_p$  [ $\text{m}^{-1}\text{t}^2$ ] denotes the compressibility of the solid skeleton, and  $\alpha_p = 1/(\lambda_l + 2\mu_l)$  holds. Deformation of the solid skeleton is taken into account by equation (2.45). Using the result of equation (2.55) the change in porosity can be related to the change in pore pressure as

$$\frac{1}{1-n} \frac{\partial n}{\partial t} = \alpha_p \frac{\partial p}{\partial t}, \quad (2.56)$$

For small time steps this equation can be expressed as  $dn = \alpha_p(1-n)dp$ , and for the porosity the following relation holds

$$n = 1 - (1 - n_0) \exp\{-\alpha_p(p - p_0)\}, \quad (2.57)$$

where  $n_0$  [-] is the reference porosity at reference pressure  $p_0$  [ $\text{ml}^{-1}\text{t}^{-2}$ ]. For small deformations its linearized form reads

$$n = n_0 + \alpha_p(1 - n_0)(p - p_0). \quad (2.58)$$

Freeze and Cherry [52] propose  $\alpha_p = 1 \cdot 10^{-10} \text{ ms}^2/\text{kg}$  for sound bedrock and  $\alpha_p = 1 \cdot 10^{-7} \text{ ms}^2/\text{kg}$  for clay. It must be noted that these expressions only hold for the one-dimensional dissipation of pore-water pressures, and do not capture pore pressure generation by mechanical or hydraulical loading on the subsurface.

## 2.2.2 Saturation

Constitutive equations for the degree of saturation of a fluid in a porous medium are based on the macroscopic capillary pressure. The capillary pressure denotes the difference of the non-wetting and the wetting phase pressure. For liquid-gas flow systems the liquid phase is the wetting phase and the gas phase is the non-wetting phase. The macroscopic capillary pressure  $p_c$  [ $\text{ml}^{-1}\text{t}^{-2}$ ] reads

$$p_c = p^g - p^l. \quad (2.59)$$

If the gas phase is stagnant, the pressure in the gas phase is constant and equals the atmospheric pressure if the phase is continuous. Then,  $p_c = -p^l$  holds for unsaturated conditions and  $p_c = 0$  applies for saturated conditions. In general the constitutive relation for two immiscible phases reads  $p_c = p_c(S^l, T^l, T^g, \omega_k^l, \omega_k^g)$ . Van Genuchten and Brooks-Corey simplified this by an empirical relation  $p_c = p_c(S)$ , where  $S$  denotes the saturation of the wetting phase [117, 9].

Reversely, Van Genuchten-Mualem express the saturation as a functional relation of wetting phase pressure according to  $S = S(p)$ . The Van Genuchten model is written as

$$\begin{cases} S = S_r + (S_s - S_r) [1 + |g_a \psi|^{g_n}]^{-g_m} & \text{if } \psi < \psi_a \\ S = S_s & \text{for } \psi \geq \psi_a \end{cases}, \quad (2.60)$$

where  $\psi$  [l] denotes the pressure head  $\psi = p^l/\rho^l g$ ,  $\psi_a$  [l] is the air-entry pressure head, which is constraint by  $\psi_a \leq 0$ .  $S_r$  [-] is the minimal saturation and  $S_s$  [-] denotes the maximum degree of saturation. The minimal saturation deviates from zero due to chemically attached water or entrapped water pockets. The maximum degree of

saturation is less than one mainly as a result of entrapped air. The Van Genuchten relation counts two empirical shape factors that have to be measured in the laboratory:  $g_n$  [-] and  $g_a$  [ $l^{-1}$ ]. For convenience a third shape factor (Mualem assumption) was introduced as  $g_m = (g_n - 1)/g_n$  [-].

Table 2.1: Staringreeks Van Genuchten shape factors.

soil type	$S_r$	$n$	$K_d$	$g_a$	$g_l$	$g_n$
	-	-	m/d	1/m	-	-
O1 sand	0.03	0.36	0.1522	2.24	0.000	2.286
O8 sandy clay	0.00	0.47	0.0908	1.36	-0.803	1.342
O11 clay	0.00	0.42	0.0138	1.91	-1.384	1.152
O14 loam	0.03	0.38	0.0151	0.30	-0.292	1.728
O16 peat	0.00	0.89	0.0107	1.03	-1.411	1.376

Figure 2.1 shows the functional saturation relation for five Dutch soil types, which are reported in the Staring series [111], and table 2.1 shows their parameters. The relative permeability relates to the saturation of the soil and will be discussed in the next section. Instead of saturation the Staring series use water content  $\theta$  [-], which relates to saturation and porosity as  $\theta = nS$ .

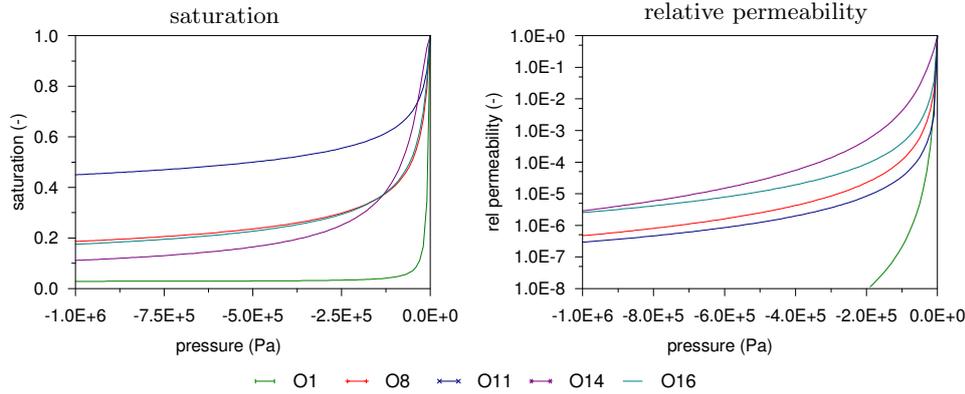


Figure 2.1: Saturation and relative permeability for Staring series materials.

The Staring series do not provide a value for the porosity and here the porous medium is assumed to be fully saturated for  $S_s = 1$ . Under this assumption the porosity follows from the maximum water content  $n = \theta_s$  and the residual saturation follows from  $S_r = \theta_r/n$ . For time dependent problems the pressure derivative of the saturation may be used. This derivative reads

$$\frac{dS}{dp} = \frac{g_a(g_n - 1)(S_s - S_r)(g_a\psi)^{(g_n-1)}}{[1 + (g_a\psi_p)^{g_n}]^{(2g_n-1)/g_n}}. \quad (2.61)$$

Alternatively Brooks-Corey parametric model is given by

$$\begin{cases} S = S_r + (S_s - S_r) |b_\alpha \psi|^{-b_n} & \text{if } \psi < -b_\alpha^{-1} \\ S = S_s & \text{for } \psi \geq -b_\alpha^{-1} \end{cases}, \quad (2.62)$$

where  $b_\alpha$  [ $l^{-1}$ ] is a curve fitting parameter,  $b_n$  [-] denotes the pore distribution index  $b_n \geq 1$ .

Many materials show a non-unique relationship for the saturation and wetting phase pressure. This hysteresis is caused by the variation in pore diameters, differences in radii of advancing and receding menisci, air entrapment etc. Van Dam [111] proposes a main wetting curve and a main drying curve. The main drying curve describes the gradual desorption of an initially saturated sample, and the main wetting curve follows from an gradual adsorption experiment of an initially dry sample. Partially drying and partially wetting follow drying scanning curves and wetting scanning curves. In general, the degree of saturation for a given pressure will be higher on the main drying curve than the saturation on the main wetting curve. Figure 2.1 shows the main drying curves for saturation as a function of pore pressure. The hysteresis can be captured by an alternative relationship, but it remains typically empirical [84, 114].

### 2.2.3 Permeability

For saturated conditions, the permeability of the porous medium is expressed by the intrinsic permeability tensor. The intrinsic permeability tensor is a second order tensor. If the working lines of  $\mathbf{K} \cdot \mathbf{n}$  coincide, then  $\mathbf{n}$  is called an eigenvector of  $\mathbf{K}$ , the corresponding eigenvalues  $\lambda_i$  follow from  $(\mathbf{K} - \lambda \mathbf{I}) \cdot \mathbf{n} = \mathbf{0}$ . The eigenvalues correspond to principal permeabilities and the eigenvectors capture the layering of the porous medium.

If the porous medium is partly saturated then the intrinsic permeability is scaled by the relative permeability, which is often given as a function of saturation. The permeability of the porous medium for a fluid phase ( $\alpha \in l, g$ ) is thus given by

$$\mathbf{K}^\alpha = k_r^\alpha (S^\alpha) \mathbf{K}, \quad (2.63)$$

where the relative permeability is denoted by  $k_r$  [-] and its value holds  $0 < k_r \leq 1$ . Mualem-Van Genuchten and Brooks-Corey proposed empirical relations for the relative permeability [117, 9]. The Van Genuchten relation reads

$$k_r = (S_e)^{g_l} \left[ 1 - \left( 1 - S_e^{1/g_m} \right)^{g_m} \right]^2, \quad S_e = \frac{S - S_r}{S_s - S_r}. \quad (2.64)$$

The empirical shape factor  $g_l$  [-] is often set to 0.5. Figure 2.1 presents the material behavior for five Dutch soils: sand O1, sandy clay O8, clay O11, loam O14 and peat O16. For these materials the Van Genuchten shape factors and the Darcy permeability were given in table 2.1. The Darcy permeability tensor  $\mathbf{K}_d$  [ $lt^{-1}$ ] is derived from the intrinsic permeability tensor  $\mathbf{K}$  and reads

$$\mathbf{K}_d = \frac{\mathbf{K} \rho g}{\mu}, \quad (2.65)$$

where  $g$  denotes the magnitude of the gravity vector. The Darcy permeability tensor expresses a bulk property of the porous medium, as the intrinsic permeability tensor depends on the solid phase, and density and dynamic viscosity are fluid properties. The Darcy permeability will not be used in this thesis since viscosity and density vary in time and space.

Brooks-Corey's parametric model expresses the relative permeability as

$$k_r = S_e^{b_\kappa}, \quad S_e = \frac{S - S_r}{S_s - S_r}, \quad (2.66)$$

where  $b_\kappa = 2/b_n + b_l + 2$  [-] is a curve fitting parameter and  $b_l$  [-] is known as the pore connectivity parameter.

## 2.2.4 Density

This subsection considers the density of a binary fluid, composed of pure water and dissolved salt in the range from pure water to saturated brine. The general equation of state for the density of this liquid phase reads

$$\begin{aligned} d\rho^\alpha &= \left( \frac{1}{\rho^\alpha} \frac{\partial \rho^\alpha}{\partial p^\alpha} \right) \rho^\alpha dp^\alpha + \left( \frac{1}{\rho^\alpha} \frac{\partial \rho^\alpha}{\partial \omega_k^\alpha} \right) \rho^\alpha d\omega_k^\alpha + \left( \frac{1}{\rho^\alpha} \frac{\partial \rho^\alpha}{\partial T^\alpha} \right) \rho^\alpha dT^\alpha \\ &= \beta_p \rho^\alpha dp^\alpha + \beta_\omega \rho^\alpha d\omega_k^\alpha - \beta_T \rho^\alpha dT^\alpha, \end{aligned} \quad (2.67)$$

where  $\beta_p$  [ $\text{m}^{-1}\text{t}^2$ ] denotes the liquid phase compressibility,  $\beta_{\omega_k}$  [-] is the volumetric solute coefficient, and  $\beta_T$  [ $\text{T}^{-1}$ ] is the thermal expansion coefficient. For constant compressibility and expansion coefficients the liquid density can be written as

$$\rho^l = \rho_0^l \exp\{\beta_p(p - p_0) + \beta_\omega(\omega - \omega_0) - \beta_T(T - T_0)\}. \quad (2.68)$$

Here  $p_0$  [ $\text{ml}^{-1}\text{t}^{-2}$ ] denotes the reference fluid pressure,  $\omega_0$  [-] is the reference salt concentration, and  $T_0$  [T] is the reference temperature. At these reference values the density of the fluid, expressed as  $\rho_0^l = \rho_0^l(p_0, \omega_0, T_0)$  [ $\text{ml}^{-3}$ ], is determined. A first order Taylor expansion gives a linear approximation, which reads

$$\rho^l = \rho_0^l [1 + \beta_p(p - p_0) + \beta_\omega(\omega - \omega_0) - \beta_T(T - T_0)]. \quad (2.69)$$

The density of the fluid depends primary upon the fluid solute concentration and its temperature. In the range of 293 K to 333 K the thermal expansion coefficient  $\beta_T$  is about  $3.75 \cdot 10^{-4}$  1/K for pure water ( $\omega_0 = 0$ ). The thermal expansion coefficient for fresh water in the range of 273 K to 373 K varies from  $-0.68 \cdot 10^{-4}$  1/K to  $7.5 \cdot 10^{-4}$  1/K and is zero at 277 K. The solute coefficient  $\beta_\omega$  is approximately 0.7 for pure water and salt mixtures, at 293 K,  $\rho_0$  equals 998.2 kg/m<sup>3</sup>. The density of seawater ( $\omega = 0.024$ ) is about 1022 kg/m<sup>3</sup>, for a saturated brine ( $\omega = 0.302$ ) the density can be as high as 1300 kg/m<sup>3</sup>. The weak dependency of density with pressure follows from the measured compressibility coefficient: for pure water at 293 K  $\beta_p$  is about  $4.47 \cdot 10^{-10}$  ms<sup>2</sup>/kg.

### 2.2.5 Viscosity

The dynamic viscosity of a binary fluid of pure water and dissolved salt constituents, can be expressed empirically [15] as

$$\mu^l = \mu_0^l (\gamma_0 + \gamma_1 \omega + \gamma_2 \omega^2 + \gamma_3 \omega^3), \quad (2.70)$$

where  $\mu_0^l$  [ml<sup>-1</sup>t<sup>-1</sup>] is the reference viscosity and  $\gamma_i$  [-] are polynomial constants. Lever and Jackson [89] determined the coefficients for dissolved salt (NaCl) as  $\gamma_0 = 1$ ,  $\gamma_1 = 1.85$ ,  $\gamma_2 = -4.1$  and  $\gamma_3 = 44.5$ . According to this relation the viscosity for seawater is about  $1.04 \cdot 10^{-3}$  kg/ms and for saturated brine it equals  $2.39 \cdot 10^{-3}$  kg/ms. The reference viscosity of pure water at 293 K equals  $1.0 \cdot 10^{-3}$  kg/ms.

The dynamic viscosity of pure water, as a function of temperature, is given by

$$\mu^l = (243.18 \cdot 10^{-7}) 10^{247.8/T-140}. \quad (2.71)$$

According to this equation the viscosity of pure water at 293 K equals  $1.01 \cdot 10^{-3}$  kg/ms, and at 333 K the viscosity is  $4.68 \cdot 10^{-3}$  kg/ms.

## 2.3 Problem definition

This section poses the mathematical model, subjected to initial and boundary conditions, which needs to be solved in order to obtain the solution for a specific subsurface flow problem. The mathematical model is composed out of the previously given balance equations and the constitutive relations, and forms a set of coupled nonlinear partial differential equations formulated in terms of independent primary variables. Physically admissible initial conditions and boundary conditions pose coupled constraints.

Section 2.3.1 states the system of equations composed by the flow equation, the mass transport equation and the heat transport equation, and selects the primary variables: fluid pressure, solute mass fraction, and porous medium temperature. These basic equations are coupled by expressions for the convective flux and dispersive fluxes, and the constitutive relations for porosity, saturation, permeability, density and viscosity. The section also presents the Rayleigh number that characterize the flow regime in the case of density driven flow. Section 2.3.2 gives a set of initial conditions and section 2.3.3 gathers boundary conditions for the coupled set of equations. This section considers flow of a binary liquid, and drops the liquid phase index and the chemical species index.

### 2.3.1 System of differential equations

The flow equation follows from the mass balance equation of the liquid phase given by equation (2.10) and Darcy's law presented by equation (2.26) as

$$n \frac{\partial \rho S}{\partial t} + S \rho \alpha_p \frac{\partial p}{\partial t} - \nabla \cdot \left( \rho \frac{k_r \mathbf{K}}{\mu} \cdot (\nabla p - \rho \mathbf{g}) \right) = \tilde{\rho} \bar{Q}. \quad (2.72)$$

The pressure in the liquid phase  $p$  is the primary unknown in the flow equation. In this formulation the compressibility of the solid skeleton  $\alpha_p$  captures the time dependent

variation of porosity according to equation (2.56). The dynamic viscosity  $\mu$ , expressed by equation (2.71), of the water phase has a functional relation with temperature. The density  $\rho$  of the water phase, according to equation (2.69), depends mainly on the mass fraction of the constituents (fresh water and dissolved salt). The porosity  $n$ , according to equation (2.58), depends on the fluid pressure  $p$  only. The intrinsic permeability  $\mathbf{K}$  of the solid skeleton is assumed to be constant in time. The Van Genuchten model poses a unique relationship of saturation  $S$  with pore pressure by equation (2.60), and relates relative permeability  $k_r$  to pore pressure, according to equation (2.64). The volumetric source term  $\overline{Q}$  in the right hand side is predefined. Alternatively a prescribed mass source  $\overline{M} = \overline{\rho} \overline{Q}$  can be applied. The density of the fluid  $\overline{\rho}$  is prescribed for positive values of the source term (injection into the subsurface). The gravitational acceleration  $\mathbf{g}$  is constant.

The solute transport equation follows from the mass transport equation for a water phase component, as given by equation (2.14). If the non-convective flux is replaced by the Fickian type dispersion equation, given by equation (2.27) and chemical reactions are neglected, then the convective form of the solute transport equation reads

$$nS\rho \frac{\partial \omega}{\partial t} + \rho \mathbf{q} \cdot \nabla \omega - \nabla \cdot \rho \mathbf{D} \cdot \nabla \omega = \tilde{\rho} (\tilde{\omega} - \omega) \overline{Q}. \quad (2.73)$$

The mass fraction  $\omega$  of the chemical species is the primary unknown in this equation. The Darcy flux  $\mathbf{q}$  couples the flow equation to the mass transport equation via the convective term and the dispersive term in the solute transport equation, since the hydrodynamic dispersion  $\mathbf{D}$  incorporates the Darcy flux. Darcy's law is given by equation (2.26), the hydrodynamic dispersion follows from equation (2.28). The right hand side cancels out for a negative value of the volumetric source (extraction out of the subsurface) where  $\tilde{\omega} = \omega$  holds.

The heat transport equation follows from the combined liquid and solid phase energy balance equation (2.41). If the non-convective heat flux is captured by equation (2.42), then the heat transport equation reads

$$(cnS\rho + (1-n)c^s\rho^s) \frac{\partial T}{\partial t} + c\rho \mathbf{q} \cdot \nabla T - \nabla \cdot \mathbf{H} \cdot \nabla T = c\tilde{\rho} (\tilde{T} - T) \overline{Q}, \quad (2.74)$$

where temperature  $T$  is the primary variable, which applies to the solid skeleton and the groundwater. The heat capacity of the water phase  $c$  is assumed to be constant. The density of the solid phase  $\rho^s$  and the specific heat capacity of the solid matrix  $c^s$  are assumed to be constant in time but may vary in space. The Darcy flux  $\mathbf{q}$  couples the flow equation to the heat transport equation. The hydrodynamic thermal dispersion  $\mathbf{H}$  incorporates the Darcy flux according to equation (2.43). For a positive value of the volumetric source term, the temperature of the injected water is predefined as  $\tilde{T} = \overline{T}$  holds. For a negative volumetric source term the right hand side of the heat transport equation becomes zero.

### Characteristic numbers

Heat and solute concentration differences may introduce potentially unstable situations if a lower density fluid lies underneath a higher density fluid [41, 121, 66]. In a sub-critical conductive system, characterized by low Rayleigh numbers, density changes

only with depth. Flow occurs when the critical Rayleigh number is exceeded. If  $Ra > 4\pi^2$ , the conduction state in a porous medium heated from below becomes unstable [66]. In this first critical state there are three solutions: two stable steady convection solutions and one unstable conduction solution. The convection solutions show convection cells, their flow direction is opposite to each other. The rotation direction depends on the initial disturbance. In bifurcation theory this is called a pitch fork bifurcation. The steady convection solutions become unstable if the Rayleigh number increases further. At the second critical Rayleigh number the flow regime changes again. At this point the steady convective regime becomes unstable and there is a transition to a fluctuating convective regime. At this point a Hopf bifurcation occurs and the steady rotating state switches to an oscillatory rotating state, for which the rotating velocity changes in time. Other critical numbers may be observed for even higher Rayleigh numbers.

Characteristic Rayleigh numbers will be derived for a simplified set of equations. At first, flow in a fully saturated non-deforming porous medium in the absence of source terms will be considered. Secondly the Oberbeck-Boussinesq approximation will be imposed. This approximation simplifies the flow equation and both transport equations by neglecting density changes, except for the buoyancy term  $\rho\mathbf{g}$  in Darcy's law. The flow equation (2.72) then follows from the mass balance equation for the fluid and Darcy's law as

$$\nabla \cdot \mathbf{q} = 0, \quad \mathbf{q} = -\frac{K}{\mu} \cdot (\nabla p - \rho\mathbf{g}). \quad (2.75)$$

The solute transport equation (2.73) simplifies to

$$n \frac{\partial \omega}{\partial t} + \mathbf{q} \cdot \nabla \omega - n \nabla \cdot \mathbf{D} \cdot \nabla \omega = 0, \quad \mathbf{D} = DI. \quad (2.76)$$

This expression imposes Fick's law, which replaces the hydrodynamic dispersion by a constant diffusion term. The heat transport equation (2.74), supplemented by Fick's law for thermal conductivity reads

$$(c\rho)^e \frac{\partial T}{\partial t} + c\rho\mathbf{q} \cdot \nabla T - (c\rho)^e \nabla \cdot \mathbf{H} \cdot \nabla T = 0, \quad \mathbf{H} = HI, \quad (2.77)$$

where  $(c\rho)^e = (nc\rho + (1-n)c^s\rho^s)$ ,  $\lambda^e = n\lambda^l + (1-n)\lambda^s$ , and  $H = \lambda^e / (c\rho)^e$ .

The Rayleigh numbers  $Ra$ , which characterize the onset of natural convection, will be derived on a two-dimension Cartesian coordinate system. First the stream function  $\Psi$  will be introduced. The definition of the stream function in a two-dimensional Cartesian domain follows from  $q_x = -\partial\psi/\partial z$  and  $q_z = \partial\psi/\partial x$  and  $d\psi = q_x dz - q_z dx$ . Continuity of mass according to equation (2.75) gives

$$-\frac{\partial^2 \Psi}{\partial z \partial x} + \frac{\partial^2 \Psi}{\partial x \partial z} = 0. \quad (2.78)$$

This relation proves that the stream function is a single-valued function [32]. Darcy's law according to equation (2.75), restated for the two-dimensional case in a homogeneous domain, reads

$$\frac{\mu}{K} q_x = -\frac{\partial p}{\partial x}, \quad \frac{\mu}{K} q_z = -\frac{\partial p}{\partial z} + \rho g. \quad (2.79)$$

Pressure is a physical scalar which has one value in every field point [32], and the following relation holds

$$-\frac{\partial^2 p}{\partial x \partial z} + \frac{\partial^2 p}{\partial z \partial x} = 0. \quad (2.80)$$

Differentiating Darcy's flux components and subtracting them gives

$$\frac{\mu}{K} \left( \frac{\partial q_x}{\partial z} - \frac{\partial q_z}{\partial x} \right) = \frac{\partial \rho g}{\partial x}. \quad (2.81)$$

Reformulated in stream functions, this relations reads

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial z^2} = -\frac{gK}{\mu} \left( \frac{\partial \rho}{\partial \omega} \frac{\partial \omega}{\partial x} + \frac{\partial \rho}{\partial T} \frac{\partial T}{\partial x} \right). \quad (2.82)$$

The linear constitutive relation (2.69) gives an expression for the density derivatives, according to  $\partial \rho / \partial \omega = \rho_0 \beta_\omega$  and  $\partial \rho / \partial T = \rho_0 \beta_T$ . The Rayleigh numbers will be formulated for a coupled flow-solute transport and a coupled flow-heat transport problem separately as the Fickian diffusion term varies per problem. The Rayleigh numbers follow from a dimensionless formulation [66]. Spatial dimensions are transformed using a typical length unit  $L$ . The dimensionless coordinates follow from  $\tilde{x} = x/L$  and  $\tilde{z} = z/L$ . The time scale for the solute transport equation transforms to  $\tilde{t} = tD/L^2$ , and the dimensionless Darcy velocity reads  $\tilde{q} = qL/(nD)$ . The solute transport equation then reads

$$\frac{\partial \tilde{\omega}}{\partial \tilde{t}} = \tilde{q}_x \frac{\partial \tilde{\omega}}{\partial \tilde{x}} + \tilde{q}_z \frac{\partial \tilde{\omega}}{\partial \tilde{z}} - \frac{\partial^2 \tilde{\omega}}{\partial \tilde{x}^2} - \frac{\partial^2 \tilde{\omega}}{\partial \tilde{z}^2}, \quad (2.83)$$

where the mass fraction is normalized by  $\tilde{\omega} = (\omega - \omega_{\min})/(\omega_{\max} - \omega_{\min})$ . The stream function follows from  $\tilde{\Psi} = \Psi/(nD)$ , and the solute transport part of equation (2.82) reads

$$\frac{\partial^2 \tilde{\Psi}}{\partial \tilde{x}^2} + \frac{\partial^2 \tilde{\Psi}}{\partial \tilde{z}^2} = -\frac{g\rho_0\beta_\omega KL}{n\mu D} \frac{\partial \tilde{\omega}}{\partial \tilde{x}}, \quad (2.84)$$

where the right hand side of the original equation was multiplied by  $L/(nD)$  in order to obtain the dimensionless formulation. The solute Rayleigh number follows from the right hand side, and reads

$$\text{Ra}^\omega = \frac{g\Delta\rho KL}{n\mu D}. \quad (2.85)$$

Here  $\Delta\rho$  is the maximum density change due to the mass fraction variation.

The time scale for the heat transport equation transforms to  $\tilde{t} = tH/L^2$ , and the dimensionless Darcy velocity reads  $\tilde{q} = qL(\rho c)/H(\rho c)^e$ . The heat transport equation then reads

$$\frac{\partial \tilde{T}}{\partial \tilde{t}} = \tilde{q}_x \frac{\partial \tilde{T}}{\partial \tilde{x}} + \tilde{q}_z \frac{\partial \tilde{T}}{\partial \tilde{z}} - \frac{\partial^2 \tilde{T}}{\partial \tilde{x}^2} - \frac{\partial^2 \tilde{T}}{\partial \tilde{z}^2}, \quad (2.86)$$

where the temperature is normalized by  $\tilde{T} = (T - T_{\min})/(T_{\max} - T_{\min})$ . The dimensionless stream function for this problem follows from  $\tilde{\Psi} = (\rho c)/(\rho c)^e \cdot \Psi/H$ , the heat transport part of equation (2.82) reads

$$\frac{\partial^2 \tilde{\Psi}}{\partial \tilde{x}^2} + \frac{\partial^2 \tilde{\Psi}}{\partial \tilde{z}^2} = \frac{g\rho c\rho_0\beta_T KL}{\mu\lambda^e} \frac{\partial \tilde{T}}{\partial \tilde{x}}, \quad (2.87)$$

where the right hand side of the original equation was multiplied by  $(\rho c L)/H(\rho c)^e$ , in order to obtain the dimensionless formulation. The thermal Rayleigh number for a linear expression of the density state reads

$$\text{Ra}^T = \frac{g\rho c \Delta\rho K L}{\mu\lambda^e}, \quad (2.88)$$

where  $\Delta\rho$  is the change in density due to the variation in temperature.

### 2.3.2 Initial conditions

Initial conditions specify the state of the system at the start of the simulation. These conditions may be specified by the distribution of the primary variables in space as

$$p = p^0 \quad \cap \quad \omega = \omega^0 \quad \cap \quad T = T^0 \quad \text{on } \Omega, \quad (2.89)$$

where  $p^0 = p(\mathbf{x}, t^0)$  is the initially prescribed pressure field,  $\omega^0 = \omega(\mathbf{x}, t^0)$  denotes the prescribed mass fraction field,  $T^0 = T(\mathbf{x}, t^0)$  is the initial temperature field, and  $\Omega$  denotes the flow domain. Alternatively, a set of derived parameter fields, like a velocity field, a density field or a viscosity field could specify the starting conditions. In general, the fields of primary variables or derived variables do not construct an equilibrium condition. Deviations from the equilibrium condition will dissipate during the transient computation. An initial state at equilibrium simplifies the analysis of simulation results. A set of boundary conditions that construct the initial state can be used to compute the initial situation at equilibrium. For calculation speed it is preferable to start with a guess of the initial state of the system that does not vary too much from the equilibrium situation.

### 2.3.3 Boundary conditions

This subsection proposes a set of physically admissible boundary conditions. Continuity of total mass flux, solute mass flux and heat energy flux requires that the fluxes on both sides of the boundary should be the same. This condition is formulated as

$$\begin{aligned} \rho\mathbf{q}\cdot\mathbf{n}|_- = \rho\mathbf{q}\cdot\mathbf{n}|_+ \quad \cap \quad (\rho\omega\mathbf{q} - \rho\mathbf{D}\cdot\nabla\omega)\cdot\mathbf{n}|_- = (\rho\omega\mathbf{q} - \rho\mathbf{D}\cdot\nabla\omega)\cdot\mathbf{n}|_+ \\ \cap \quad (c\rho\mathbf{q}T - \mathbf{H}\cdot\nabla T)\cdot\mathbf{n}|_- = (c\rho\mathbf{q}T - \mathbf{H}\cdot\nabla T)\cdot\mathbf{n}|_+ \quad \text{on } \Gamma, \end{aligned} \quad (2.90)$$

where  $\mathbf{n}$  is the outward normal,  $|_-$  indicates the flux inside the domain and  $|_+$  denotes the flux outside the flow domain over the boundary  $\Gamma$ . For an in-flow boundary condition from a well mixed reservoir (outside the porous medium) the diffusive fluxes just outside the flow domain can be neglected ( $\rho\mathbf{D}\cdot\nabla\omega)\cdot\mathbf{n}|_+ \approx 0$  and  $(\mathbf{H}\cdot\nabla T)\cdot\mathbf{n}|_+ \approx 0$ ).

Equation (2.90) applies directly to the divergent form of the transport equations. A more convenient formulation of the boundary conditions for the convective form of the transport equations reads

$$\begin{aligned} -\rho\mathbf{q}\cdot\mathbf{n} = \tilde{\rho}\tilde{q}_n \quad \cap \quad (\rho\mathbf{D}\cdot\nabla\omega)\cdot\mathbf{n} = \tilde{\rho}\tilde{q}_n(\tilde{\omega} - \omega) \\ \cap \quad (\mathbf{H}\cdot\nabla T)\cdot\mathbf{n} = c\tilde{\rho}\tilde{q}_n(\tilde{T} - T) \quad \text{on } \Gamma. \end{aligned} \quad (2.91)$$

This condition replaces the convective flux outside the porous medium by the product of a prescribed volumetric influx  $\bar{q}_n$  and a variable liquid phase density  $\bar{\rho}$ , which equals the density of the liquid phase within the medium  $\rho$  if outflow occurs or is set to a prescribed density  $\bar{\rho}$  if inflow takes place. The convective flux reads  $\rho \mathbf{q} \cdot \mathbf{n}|_+ = -\bar{\rho} \bar{q}_n$ . According to equation (2.91), the advective species mass flux equals  $\bar{\rho} \bar{q}_n \bar{\omega}$  for flow into the medium given by positive values of  $\bar{q}_n$ . For outflow the advective mass flux depends on the density and mass fraction inside the porous medium and reads  $-\rho \bar{q}_n \omega$ . The diffusive flux for this condition is set to zero, as  $\tilde{\omega} = \omega$ . The advective heat flux reads  $c \bar{\rho} \bar{q}_n \bar{T}$  for inflow and  $c \rho \bar{q}_n T$  for outflow. The average porous medium diffusive heat flux is set to zero for outflow conditions.

For no-flow boundary conditions the total mass flux, the solute mass flux, and the heat energy flux across the boundary are zero. If the convective flux is set to zero then the diffusive solute mass flux and diffusive energy flux are zero as well. This boundary condition reads

$$\rho \mathbf{q} \cdot \mathbf{n} = 0 \quad \cap \quad (\rho \mathbf{D} \cdot \nabla \omega) \cdot \mathbf{n} = 0 \quad \cap \quad (\mathbf{H} \cdot \nabla T) \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_1. \quad (2.92)$$

These boundary conditions for the flow equation and for the transport equations are known as a Neumann condition.

To be consistent with the source term  $\bar{Q}$  in the governing equations (2.72), (2.73) and (2.74), the prescribed boundary flux  $\bar{q}_n$  will be positive for flow into the porous medium and negative for flow out of the porous medium. For  $\bar{q}_n > 0$  the boundary condition reads

$$\begin{aligned} -\rho \mathbf{q} \cdot \mathbf{n} = \bar{\rho} \bar{q}_n \quad \cap \quad (\rho \mathbf{D} \cdot \nabla \omega) \cdot \mathbf{n} = \bar{\rho} \bar{q}_n (\bar{\omega} - \omega) \\ \cap \quad (\mathbf{H} \cdot \nabla T) \cdot \mathbf{n} = c \bar{\rho} \bar{q}_n (\bar{T} - T) \quad \text{on } \Gamma_2, \end{aligned} \quad (2.93)$$

where  $\bar{\rho} \bar{q}_n$  denotes the prescribed total mass flux,  $\bar{\rho} \bar{q}_n \bar{\omega}$  is the prescribed salt mass fraction flux, and  $c \bar{\rho} \bar{q}_n \bar{T}$  denotes the pre-defined heat energy flux. As  $\mathbf{n}$  points outside the flow domain, the in-product  $\mathbf{q} \cdot \mathbf{n}$  specifies outflow. The no-flow condition given by equation (2.92) forms a special case of this inflow condition. The constraints on the transport equations are mixed type conditions, also called Robin boundary conditions.

If a Dirichlet condition is imposed on the flow equation (prescribed pressure), then the corresponding flux can be calculated. If  $\mathbf{q} \cdot \mathbf{n} < 0$ , inflow occurs and the combined conditions read

$$\begin{aligned} p = \bar{p} \quad \cap \quad (\rho \mathbf{D} \cdot \nabla \omega) \cdot \mathbf{n} = -\bar{\rho} (\bar{\omega} - \omega) \mathbf{q} \cdot \mathbf{n} \\ \cap \quad (\mathbf{H} \cdot \nabla T) \cdot \mathbf{n} = -c \bar{\rho} (\bar{T} - T) \mathbf{q} \cdot \mathbf{n} \quad \text{on } \Gamma_3, \end{aligned} \quad (2.94)$$

where  $\bar{p}$  denotes the prescribed boundary pressure.

If the (negative) prescribed volumetric flux  $\bar{q}_n$  specifies outflow, then its value is given by the mass fraction inside the domain. This also holds for the temperature. If  $\bar{q}_n < 0$  holds then the outflow boundary condition reads

$$-\rho \mathbf{q} \cdot \mathbf{n} = \rho \bar{q}_n \quad \cap \quad (\rho \mathbf{D} \cdot \nabla \omega) \cdot \mathbf{n} = 0 \quad \cap \quad (\mathbf{H} \cdot \nabla T) \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_4. \quad (2.95)$$

This condition follows from equation (2.91) with  $\tilde{\omega} = \omega$  and  $\tilde{T} = T$ . The no-flow condition given by equation (2.92) also forms a special case of this outflow condition.

If a Dirichlet condition is imposed on the flow equation and  $\mathbf{q} \cdot \mathbf{n} > 0$  applies then the outflow condition is reformulated as

$$p = \bar{p} \quad \cap \quad (\rho \mathbf{D} \cdot \nabla \omega) \cdot \mathbf{n} = 0 \quad \cap \quad (\mathbf{H} \cdot \nabla T) \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_5. \quad (2.96)$$

For a boundary that separates two parts of the porous medium the mass-fraction and the temperature should be the same on both parts of the boundary. Inflow and outflow boundary conditions now read

$$-\rho \mathbf{q} \cdot \mathbf{n} = \bar{\rho} \bar{q}_n \quad \cap \quad \omega = \bar{\omega} \quad \cap \quad T = \bar{T} \quad \text{on } \Gamma_6. \quad (2.97)$$

Alternatively for a given pressure at the internal boundary the condition is reformulated as

$$p = \bar{p} \quad \cap \quad \omega = \bar{\omega} \quad \cap \quad T = \bar{T} \quad \text{on } \Gamma_7. \quad (2.98)$$

The previous boundary conditions impose either a pressure or a total mass flux on the flow equation. A seepage condition or free-outflow condition imposes a geometrically non-linear boundary conditions. According to this condition, the atmospheric pressure is imposed as long as outflow takes place and a no-flow condition holds if the pore pressure at the boundary (inside the flow domain) is negative. This condition is written as

$$\begin{cases} p = 0 & \text{if } \rho \mathbf{q} \cdot \mathbf{n} \geq 0 \\ \rho \mathbf{q} \cdot \mathbf{n} = 0 & \text{for } p < 0 \end{cases}. \quad (2.99)$$

The seepage condition prevents water from flowing into the domain. If water is available by precipitation, the seepage condition is generalized as the ponding condition. For this condition the prescribed influx at the boundary will be positive ( $\bar{q}_n > 0$ ) and water is allowed to flow into the domain. The rate at which inflow takes place depends on the capacity of the soil. If the pressure at the boundary exceeds the ponding pressure ( $\bar{p}_p > 0$ ) then infiltration will take place at a rate lower than the precipitation rate. Under this condition the ponding pressure prescribes the infiltration rate. The ponding pressure corresponds to the depth of small depressions in the terrain. If these depressions are filled, overland flow will occur. The ponding condition generalizes the seepage condition, by replacing the atmospheric pressure by the ponding pressure which seems realistic for outflow conditions as well, because depressions can get filled by out-flowing water.

$$\begin{cases} p = \bar{p}_p & \text{if } -\rho \mathbf{q} \cdot \mathbf{n} \leq \bar{\rho} \bar{q}_n \\ -\rho \mathbf{q} \cdot \mathbf{n} = \bar{\rho} \bar{q}_n & \text{for } p < \bar{p}_p \end{cases}. \quad (2.100)$$

The evaporation (of bare soil) boundary condition ( $\bar{q}_n < 0$ ) also generalizes the seepage boundary condition. The evaporation flux is limited by an evaporation pressure ( $\bar{p}_e < 0$ ) which follows from the humidity of air. This boundary condition is written as

$$\begin{cases} p = \bar{p}_e & \text{if } \rho \mathbf{q} \cdot \mathbf{n} \leq -\bar{\rho} \bar{q}_n \\ \rho \mathbf{q} \cdot \mathbf{n} = -\bar{\rho} \bar{q}_n & \text{for } p > \bar{p}_e \end{cases}. \quad (2.101)$$

The evaporation boundary condition limits the outflux to the flux that follows from the prescribed evaporation pressure. During evaporation, the saturation of the soil and its relative permeability decreases, and the soil becomes less capable to support the outflux. The modeling of transpiration by vegetation-roots follows the same approach, the outflux however is simulated by a volumetric source term. Compared to the bare soil situation where the flux is imposed on the boundary, the permeability of the soil will decrease less and can support the imposed outflux longer.

Seepage conditions, ponding conditions and evaporation conditions are generalized as

$$\left\{ \begin{array}{ll} p = \bar{p}_p & \text{if } \rho \mathbf{q} \cdot \mathbf{n} \geq -\rho \bar{q}_n \quad \cap \quad \bar{q}_n \geq 0 \\ -\rho \mathbf{q} \cdot \mathbf{n} = \bar{\rho} \bar{q}_n & \text{for } \bar{p}_e < p < \bar{p}_p \\ p = \bar{p}_e & \text{if } \rho \mathbf{q} \cdot \mathbf{n} \leq -\rho \bar{q}_n \quad \cap \quad \bar{q}_n < 0 \end{array} \right. . \quad (2.102)$$

According to this expression the pressure at the boundary of the flow domain is bounded between a positive pressure at ponding conditions  $\bar{p}_p$  and a negative pressure during evapotranspiration  $\bar{p}_e$ . Ponding conditions take place if the potential influx exceeds the influx capacity of the soil. This condition is enforced if the outflux found for the imposed pressure (which will have a negative value during infiltration) exceeds the negative value of the potential influx  $\bar{q}_n$ . If the ponding condition applies, the pressure reads  $p = \bar{p}_p$ , and equation (2.94) holds for liquid phase mass fraction and porous medium temperature. If the maximum evaporation condition, given by  $p = \bar{p}_e$ , holds then equation (2.96) is imposed. If the flux is prescribed by  $-\rho \mathbf{q} \cdot \mathbf{n} = \bar{\rho} \bar{q}_n$  then equation (2.93) applies for inflow and equation (2.95) holds for outflow. For embankments a free water surface could generate an extra condition on the pressure. If (part of) the boundary submerges the conditions of equation (2.102) are overruled, and infiltration or exfiltration takes place at a pre-described pressure.



# Chapter 3

## Numerical model

Analytical solutions may exist for a set of simplified flow and transport equations posed on specific homogeneous geometries and closed by linear boundary conditions. In general however, due to the nonlinearities and the coupling of the equations no analytical solutions can be found and one has to rely on approximate numerical solution methods like the finite difference method, the finite volume method or the finite element method [71, 72, 125, 13] to solve the equations. For geotechnical simulations the finite element method is preferred, as this method captures complex geometries more flexibly [116, 20, 19]. Finite element formulations are based either on a weighted residual approach or a variational approach. The weighted residual approach applies to a wider range of applications than the variational approach. However, if an equivalent minimization problem exists, both approaches produce the same set of equations. For transient boundary value problems, the method of lines transforms the partial differential equations into (nonlinear) algebraic equations. According to this method a semi-discretization in space first translates the partial differential equations into ordinary differential equations. Next a time integration scheme produces the final set of algebraic equations. Nonlinear solvers, based on Picard techniques or Newton-Raphson techniques, resolve the nonlinearities. The final set of linearized equations then can be solved by iterative solvers like conjugate gradient methods or multigrid methods. In this thesis the Galerkin weighted residual method discretizes the partial differential equations in space. A weighted finite difference method integrates the resulting ordinary differential equations in time and provides the set of algebraic equations.

Appendix B outlines the finite element method. Section 3.1 presents the numerical formulation for each of the basic equations separately. Section 3.2 gives the overall numerical problem. The problem includes the resulting set of algebraic equations for the flow problem, the solute transport problem and the heat transport problem into a single set of algebraic equations and poses initial conditions and physically admissible boundary conditions. Throughout this chapter the summation notation convention is applied and results are presented for Cartesian coordinates.

### 3.1 Basic equations

For all basic processes, a partial differential equation poses the corresponding problem. This partial differential equation needs to be supplemented by a specification of material behavior, a statement of initial conditions, and a definition of boundary conditions. The method of lines first discretizes the differential equations in space. The spatial discretization procedure includes the weak formulation of the problem and the application of Green's theorem. The spatial discretization generates a local set of ordinary differential equations. Gaussian quadrature replaces integration by summation. The method of lines then integrates the set of local ordinary differential equations in time. A finite difference time integration scheme obtains the global the set of algebraic equations.

Section 3.2.1 presents the numerical formulation of the flow equation, section 3.2.2 presents the algebraic solute transport equation, and section 3.2.3 discretizes the heat transport equation. The basic equations share their primary variables: pressure, mass fraction and temperature.

#### 3.1.1 Flow equation

The primary unknown in the flow equations is the pore pressure  $p = p(\mathbf{x}, t)$ , which is a function of the independent ordinates space and time. The trial function  $\hat{p}$  approximates the pressure field as

$$p(\mathbf{x}, t) \approx \hat{p}(\mathbf{x}, t) = p_b(t)N_b(\mathbf{x}), \quad b = 1, \dots, n_n. \quad (3.1)$$

Here  $N_b$  denotes the basis function or shape function associated with node  $b$  that interpolates the discrete nodal values of the pressure  $p_b$  in space. The number of nodes is given by  $n_n$ .

#### Problem definition

The differential equation, which describes flow of a single fluid in a porous medium given by equation (2.72) reformulated in Cartesian coordinates reads

$$n \frac{\partial \rho S}{\partial t} + \alpha_p \rho S \frac{\partial p}{\partial t} = \frac{\partial}{\partial x_i} \left[ \frac{\rho k_r K_{ij}}{\mu} \left( \frac{\partial p}{\partial x_j} - \rho g_j \right) \right] + \tilde{\rho} \bar{Q} \quad \text{on } \Omega. \quad (3.2)$$

The compressibility of the solid skeleton  $\alpha_p = \alpha_p(\mathbf{x})$ , and the components of the intrinsic permeability tensor  $K_{ij} = K_{ij}(\mathbf{x})$  are solid phase properties. Both properties are assumed to be constant in time but may vary in space. Expressions for porosity  $n$ , saturation  $S$ , relative permeability  $k_r$ , fluid density and fluid viscosity will be given at the end of this section. The gravitational acceleration vector  $[g] = [0, 0, -9.81]^T$  is constant and its components are given by  $g_i$ . The volumetric source term  $\bar{Q} = \bar{Q}(\mathbf{x}, t)$  [ $\text{t}^{-1}$ ] is predefined, and applies over the flow domain. The liquid phase density in the right hand side  $\tilde{\rho} = \bar{\rho}$  is predefined for positive values of the volumetric source (injection into the subsurface). For negative values of the volumetric source the fluid phase density  $\tilde{\rho} = \rho$  depends on the density in the porous medium (extraction from

the subsurface). An injection well located at a nodal position  $\mathbf{x}_a$  can be expressed in terms of the volumetric source term as  $Q_\delta = Q(\delta(\mathbf{x} - \mathbf{x}_a), t)$  [ $\text{m}^3\text{t}^{-1}$ ], where  $\delta$  denotes the Dirac delta function.

Initial conditions are given by a primary variable field, the pressure field  $p^0 = p(\mathbf{x}, 0)$  on the flow domain reads

$$p = p^0 \quad \text{on } \Omega. \quad (3.3)$$

Alternatively a saturation field could be given on parts of the flow domain. The functional relation only holds for the unsaturated part of the domain. However, for the saturated part of the domain, no unique relation with the pore pressure exists and for low saturation, the determination of the pressure becomes less accurate.

Three types of boundary conditions complete the flow equation. Dirichlet or first-type boundary conditions prescribe the pressure on parts of the boundary. Dirichlet boundary conditions can be expressed by

$$p = \bar{p} \quad \text{on } \Gamma_1^p, \quad (3.4)$$

where  $\Gamma_1^p$  is the part of the boundary where these conditions apply.

Von Neumann or second-type boundary conditions prescribe the derivative of the pressure or flux on the boundary. These second-type boundary conditions read

$$\frac{\rho k K_{ij}}{\mu} \left( \frac{\partial p}{\partial x_j} - \rho g_j \right) n_i = \tilde{\rho} \bar{q}_n \quad \text{on } \Gamma_2^p, \quad (3.5)$$

here  $\tilde{\rho} \bar{q}_n$  denotes the convective mass flux over the boundary, and  $\bar{q}_n$  captures the volumetric part. If  $\bar{q}_n > 0$  then flow takes place into the flow domain. The mass flux follows from  $\tilde{\rho} \bar{q}_n = \rho \bar{q}_n$ . Here  $\bar{p}$  is the known density of the in-flowing fluid. If outflow occurs then the density reads  $\tilde{\rho} = \rho$ .

Robin or third-type boundary conditions relate a flux to a pressure, and can be written as

$$\frac{\rho k K_{ij}}{\mu} \left( \frac{\partial p}{\partial x_j} - \rho g_j \right) n_i = r(\bar{p} - p) \quad \text{on } \Gamma_3^p. \quad (3.6)$$

This type of boundary condition simulates the behavior of a drain, for example. Flow out of the drain occurs if the pressure at the boundary (in the flow domain) is lower than the pressure in the drain  $\bar{p} > p$ . Outflow (into the drain) occurs in the opposite case. For this case the drain must be able to support the flux. In equation (3.6),  $r$  denotes an empirical resistance factor.

### Spatial discretization

A weak form of equation (3.2) follows from integration over the flow domain and multiplication by weighting functions  $W_a$ , which are attached to nodes. The first weak form is given by

$$\begin{aligned} \int_{\Omega^e} W_a n \frac{\partial}{\partial t} (S\rho) d\Omega + \int_{\Omega^e} W_a \alpha_p \rho S \frac{\partial \hat{p}}{\partial t} d\Omega \\ = \int_{\Omega^e} W_a \frac{\partial}{\partial x_i} \left[ \frac{\rho k K_{ij}}{\mu} \left( \frac{\partial \hat{p}}{\partial x_j} - \rho g_j \right) \right] d\Omega + \int_{\Omega^e} W_a \tilde{\rho} \bar{Q} d\Omega. \end{aligned} \quad (3.7)$$

This finite element formulation of the flow equation is given for a single element, where  $a = 1, \dots, n_{en}$  and  $n_{en}$  is the number of nodes in the element. Application of Green's theorem to the first term on the right hand side of equation (3.7) gives

$$\begin{aligned} \int_{\Omega^e} W_a n \frac{\partial}{\partial t} (S\rho) d\Omega + \int_{\Omega^e} W_a \alpha_p \rho S \frac{\partial \hat{p}}{\partial t} d\Omega &= - \int_{\Omega^e} \frac{\rho k K_{ij}}{\mu} \frac{\partial W_a}{\partial x_i} \left( \frac{\partial \hat{p}}{\partial x_j} - \rho g_j \right) d\Omega \\ &+ \int_{\Gamma^e} W_a \frac{\rho k K_{ij}}{\mu} \left( \frac{\partial \hat{p}}{\partial x_j} - \rho g_j \right) n_i d\Gamma + \int_{\Omega^e} W_a \bar{\rho} \bar{Q} d\Omega. \end{aligned} \quad (3.8)$$

Green's theorem reduces the order of integration, which makes this second form more suitable for further elaboration. Essential boundary conditions, which prescribe the primary unknown at the boundary are taken into account by a modification of the weighting function according to  $W_a = \bar{p}_a W_a$ . Substitution of the interpolation functions  $N_b$  which span the trial functions for the pressure, and replacing the weighting functions by the basis functions  $N_a$  according to Galerkin's method, results

$$\begin{aligned} \int_{\Omega^e} N_a n \frac{d}{dt} (S\rho) d\Omega + \int_{\Omega^e} N_a \alpha_p \rho S \frac{d\hat{p}}{dt} d\Omega \\ = - \int_{\Omega^e} \frac{\rho k K_{ij}}{\mu} \frac{\partial N_a}{\partial x_i} \frac{\partial N_b}{\partial x_j} p_b d\Omega + \int_{\Omega^e} \frac{\rho k K_{ij}}{\mu} \frac{\partial N_a}{\partial x_i} \rho g_j d\Omega \\ + \int_{\Gamma_2^e} N_a N_b (\bar{\rho} \bar{q}_n)_b d\Gamma + \int_{\Gamma_3^e} N_a N_b r (\bar{p} - p)_b d\Gamma + \int_{\Omega^e} N_a N_b (\bar{\rho} \bar{Q})_b d\Omega, \end{aligned} \quad (3.9)$$

where second-type and third-type boundary conditions are incorporated. Prescribed boundary fluxes and volumetric fluxes are given at the nodal points. Their values are integrated over the element boundaries and over the elements, respectively.

### Integration in time

The finite difference method translates the set of ordinary differential equations into a set of algebraic equations. The time derivative term in the flow equation is approximated tangentially as

$$\begin{aligned} \frac{d}{dt} (S\rho) \approx \frac{(S\rho)|^{p^{n+1}} - (S\rho)|^{p^n}}{\hat{p}^{n+1} - \hat{p}^n} \frac{d\hat{p}}{dt} + \frac{(S\rho)|^{\omega^{n+1}} - (S\rho)|^{\omega^n}}{\hat{\omega}^{n+1} - \hat{\omega}^n} \frac{d\hat{\omega}}{dt} \\ + \frac{(S\rho)|^{T^{n+1}} - (S\rho)|^{T^n}}{\hat{T}^{n+1} - \hat{T}^n} \frac{d\hat{T}}{dt}, \end{aligned} \quad (3.10)$$

where  $\hat{p}^{n+1}$  denotes the pressure trial function at the new time step and  $\hat{p}^n$  is the pressure trial function at the old time step. The denoted product of density and saturation  $(S\rho)|^{p^{n+1}}$  is short for  $(S\rho)|^{p^{n+1}, \omega^{n+\theta}, T^{n+\theta}}$  and its value is calculated for the pressure at the new time step; the mass fraction and the temperature are obtained for an intermediate time step  $n + \theta$  where  $0 \leq \theta \leq 1$ . The same short notation was used for  $(S\rho)|^{\omega^{n+1}}$  where the mass fraction is calculated for the new time step and pressure and temperature follow from an intermediate time step.

A global set of ordinary differential equations follows from the assemblage of the local results obtained by the spatial discretization step. The global set is written as

$$M_{ab} \frac{dp_b}{dt} = S_{ab} p_b + F_a, \quad (3.11)$$

where the capacity matrix components  $M_{ab}$  [ $l^{n_d-2}t^2$ ] are calculated as

$$M_{ab} = \bigwedge_{e=1}^{n_e} \int_{\Omega^e} N_a N_b \left( n \frac{(S\rho)^{|p^{n+1}} - (S\rho)^{|p^n}}{p^{n+1} - p^n} + \alpha_p \rho S \right) d\Omega, \quad (3.12)$$

and approximated by

$$M_{ab} \approx \bigwedge_{e=1}^{n_e} w_m^e \left[ N_a N_b \left( n \frac{(S\rho)^{|p^{n+1}} - (S\rho)^{|p^n}}{p^{n+1} - p^n} + \alpha_p \rho S \right) |J| \right]_m^e. \quad (3.13)$$

Numerical integration approximates the integral by a summation over the integration points  $m$ ,  $w_m$  are the integration weights in an isoparametric local element and  $|J|$  is the determinant of the Jacobian matrix which translates the local element position to the global element position. The number of elements is denoted by  $n_e$ . The components of the conductivity matrix  $S_{ab}$  [ $l^{n_d-2}t$ ] follow from

$$S_{ab} = - \bigwedge_{e=1}^{n_e} \int_{\Omega^e} \frac{\rho k_r K_{ij}}{\mu} \frac{\partial N_a}{\partial x_i} \frac{\partial N_b}{\partial x_j} d\Omega - \bigwedge_{e=1}^{n_e} \int_{\Gamma_3^e} N_a N_b r d\Gamma, \quad (3.14)$$

or

$$S_{ab} \approx - \bigwedge_{e=1}^{n_e} w_m^e \left( \frac{\rho k_r K_{ij}}{\mu} \frac{\partial N_a}{\partial \xi_k} \frac{\partial N_b}{\partial x_i} \frac{\partial \xi_l}{\partial x_j} |J| \right)_m^e - \bigwedge_{e=1}^{n_{b3}} w_m (r N_a N_b |J|)_m^e. \quad (3.15)$$

Here  $n_{b3}$  expresses the number of nodes on the boundary of the element where third-type boundary conditions apply. Part of the time derivative term is transferred into the force vector. The components of the force vector  $F_a$  [ $ml^{n_d-3}t^{-1}$ ] follow from

$$\begin{aligned} F_a = & \bigwedge_{e=1}^{n_e} \int_{\Omega^e} \frac{\rho k_r K_{ij}}{\mu} \frac{\partial N_a}{\partial x_i} \rho g_j d\Omega + \bigwedge_{e=1}^{n_e} \int_{\Gamma_2^e} N_a N_b (\bar{\rho} \bar{q}_n)_b d\Gamma \\ & + \bigwedge_{e=1}^{n_e} \int_{\Gamma_3^e} N_a N_b r \bar{p}_b d\Gamma + \bigwedge_{e=1}^{n_e} \int_{\Omega^e} N_a N_b (\bar{\rho} \bar{Q})_b d\Omega \\ & - \bigwedge_{e=1}^{n_e} \int_{\Omega^e} \frac{1}{\Delta t} N_a N_b \left( n (S\rho)^{|\omega^{n+1}, T^{n+1}} - (S\rho)^{|\omega^n, T^n} \right)_b d\Omega, \end{aligned} \quad (3.16)$$

local-global mapping and numerical integration gives

$$\begin{aligned} F_a \approx & \bigwedge_{e=1}^{n_e} w_m^e \left( \frac{\rho k_r K_{ij}}{\mu} \frac{\partial N_a}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \rho g_j |J| \right)_m^e + \bigwedge_{e=1}^{n_{b2}} w_m (N_a N_b |J|)_m^e (\bar{\rho} \bar{q}_n)_b \\ & + \bigwedge_{e=1}^{n_{b3}} w_m (r N_a N_b |J|)_m^e \bar{p}_b + (\bar{\rho} \bar{Q}_\delta)_a \\ & - \bigwedge_{e=1}^{n_e} w_m^e \frac{1}{\Delta t} \left[ n N_a \left( (S\rho)^{|\omega^{n+1}, T^{n+1}} - (S\rho)^{|\omega^n, T^n} \right) |J| \right]_m^e. \end{aligned} \quad (3.17)$$

Here  $n_{b2}$  and  $n_{b3}$  denote the element nodes on second-type and third-type element boundary respectively. Volumetric sources are assumed to be collocated in the nodes of the finite element grid, and simulate point sources. Boundary fluxes are integrated over the element boundary.

The  $\theta$ -weighted time integration procedure provides a first order representation of the time derivative. The linear integration scheme transfers the set of ordinary differential equations to a set of algebraic equations, and rewrites equation (3.11) as

$$M_{ab}^{n+\theta} \frac{p_b^{n+1} - p_b^n}{\Delta t} = \theta S_{ab}^{n+1} p_b^{n+1} + (1 - \theta) S_{ab}^n p_b^n + \theta F_a^{n+1} + (1 - \theta) F_a^n. \quad (3.18)$$

Both  $M_{ab}^{n+\theta}$ ,  $S_{ab}^{n+1}$  and  $F_a^{n+1}$  depend on the yet unknown solution at the new time step  $n + 1$ . A successive substitution solution procedure resolves the nonlinearities by

$$\begin{aligned} & \left( M_{ab}^{n+\theta} - \theta \Delta t S_{ab}^{n+1,r} \right) p_b^{n+1,r+1} \\ & = \left( M_{ab}^{n+\theta} + (1 - \theta) \Delta t S_{ab}^{n,r} \right) p_b^n + \theta \Delta t F_a^{n+1,r} + (1 - \theta) \Delta t F_a^n, \end{aligned} \quad (3.19)$$

where  $r$  denotes the current iteration and  $r + 1$  expresses the new (nonlinear) iteration.

### Algebraic equations

For a fully-implicit ( $\theta = 1$ ) time integration scheme equation (3.19) can be written as

$$\left( M_{ab}^{n+1,r} - \Delta t S_{ab}^{n+1,r} \right) p_b^{n+1,r+1} = M_{ab} p_b^n + \Delta t F_a^{n+1,r}. \quad (3.20)$$

The final set of linearized algebraic equations, which approximates the flow equation then reads

$$A_{ab}^p p_b = B_a^p. \quad (3.21)$$

Here  $p_b$  is short for  $p_b^{n+1,r+1}$ . The components of the stiffness matrix  $A_{ab}^p$  [ $\text{m}^{\text{d}-2}\text{t}^2$ ] read

$$\begin{aligned} A_{ab}^p = & \sum_{e=1}^{n_e} w_m^e (N_a N_b N_c |J|)_m^e \left( n^{n+1} \frac{(S\rho)^{p^{n+1}} - (S\rho)^{p^n}}{p^{n+1} - p^n} + \alpha_p (\rho S)^{n+1} \right)_c \\ & + \Delta t \sum_{e=1}^{n_e} w_m^e \left( K_{ij} N_c \frac{\partial N_a}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_b}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} |J| \right)_m^e \left( \frac{\rho k}{\mu} \right)_c^{n+1} \\ & + \Delta t \sum_{e=1}^{n_{b3}} w_m^e (r^{n+1} N_a N_b |J|)_m^e. \end{aligned} \quad (3.22)$$

Row-sum lumping improves the condition of the matrix and generates an explicit formulation for the pressure when an Euler forward time integration scheme is adopted. Replacing  $N_b$  by  $\delta_{ab}$  produces a diagonal capacity matrix. However, the accuracy of the solution decreases by lumping the matrix. For this reason the conductivity matrix is not altered by a lumping procedure. The components of the force vector  $B_a^p$  [ $\text{m}^{\text{d}-3}$ ]

read

$$\begin{aligned}
B_a^p &= \prod_{e=1}^{n_e} w_m^e (N_a N_b N_c |J|)_m^e \left( n^{n+1} \frac{(S\rho)^{p^{n+1}} - (S\rho)^{p^n}}{p^{n+1} - p^n} + \alpha_p \rho S \right)_c p_b^n \\
&+ \Delta t \prod_{e=1}^{n_e} w_m^e \left( K_{ij} N_c \frac{\partial N_a}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \rho g_j |J| \right)_m^e \left( \frac{\rho k}{\mu} \right)_c^{n+1} + \Delta t \prod_{e=1}^{n_{b2}} w_m (N_a N_b |J|)_m^e (\tilde{\rho} \bar{q}_n)_b^{n+1} \\
&\quad + \Delta t \prod_{e=1}^{n_{b3}} w_m (r^{n+1} N_a N_b |J|)_m^e \bar{p}_b^{n+1} + \Delta t (\tilde{\rho} \bar{Q}_\delta)_a^{n+1} \\
&- \prod_{e=1}^{n_e} w_m^e (N_a N_b |J|)_m^e \left( n^{n+1} (S\rho)^{\omega^{n+1}, T^{n+1}} - n^{n+1} (S\rho)^{\omega^n, T^n} \right)_b. \quad (3.23)
\end{aligned}$$

According to this equation porosity, saturation, relative permeability, fluid density and fluid viscosity need to be calculated at the nodes. Only their products are interpolated over the elements. This approach reduces the order of the functions on the element and the algorithm does not require higher order integration schemes. It reduces the number of integration points and saves computational time needed to set up the equations.

The nodal point porosity  $(n)_a$  follows from equation (2.58) as

$$(n)_a = n_0 + \alpha_p (1 - n_0) (p_a - p_0), \quad (3.24)$$

where  $n_0$  and  $p_0$  are constant over the element. Because adjacent elements may have different material parameters the porosity field shows discontinuities over element boundaries which coincide with material domain boundaries. This is also the case for synthetic fields, where the porosity is attached to the elements rather than the nodes. The porosity in integration point  $m$  follows from interpolation  $n_m = N_a^m n_a$ .

The degree of saturation at the nodal points  $S_a$  follows from equation (2.60) as

$$\begin{cases} (S)_a = S_r + (S_s - S_r) [1 + |g_a \psi|^{g_n}]^{-g_m} & \text{if } \psi < 0 \\ (S)_a = S_s & \text{for } \psi \geq 0 \end{cases}, \quad (3.25)$$

where  $\psi = |p_a/(\rho g)|$ . The empirical shape factors  $g_a$ ,  $g_n$  and  $g_m = (g_n - 1)/g_n$ , the minimal degree of saturation  $S_r$  and the maximum degree of saturation  $S_s$  are constant over the element. The saturation in integration point  $m$  follows from interpolation  $S_m = N_a^m S_a$ . Although the pressure field is continuous, the saturation field shows discontinuities for (partly) homogeneous situations.

The nodal values of the relative permeability  $(k_r)_a$  according to equation (2.64) follows from the degree of saturation as

$$\begin{cases} (k_r)_a = (S_e)^{g_l} \left[ 1 - \left( 1 - S_e^{1/g_m} \right)^{g_m} \right]^2 & \text{if } \psi < 0 \\ (k_r)_a = 1 & \text{for } \psi \geq 0 \end{cases}, \quad (3.26)$$

where  $S_e = (S_a - S_r)/(S_s - S_r)$ . The shape factor  $g_l$  is constant over the element, the relative permeability in the integration points follows from interpolation  $(k_r)_m = N_a^m (k_r)_a$ . Like the saturation field the relative permeability field may show discontinuities.

The nodal point density  $(\rho)_a$  follows from equation (2.69) as

$$(\rho)_a = \rho_0 [1 + \beta_p(p_a - p_0) + \beta_\omega(\omega_a - \omega_0) - \beta_T(T_a - T_0)]. \quad (3.27)$$

The reference values  $\rho_0$ ,  $p_0$ ,  $\omega_0$ , and  $T_0$  are constant, as well as the coefficients  $\beta_p$ ,  $\beta_\omega$  and  $\beta_T$ . The density at the integration point follows from  $\rho^m = N_a^m \rho_a$ . The density field is  $C_0$  continuous for linear interpolation functions.

The dynamic viscosity of the fluid phase at the nodal points  $(\mu)_a$ , follows from the nodal point temperature, and is written according to equation (2.71) as

$$(\mu)_a = (2.4318 \cdot 10^{-5}) 10^{247.8/T_a - 140}. \quad (3.28)$$

Multiplying the nodal values by the interpolation functions gives a continuous viscosity field. At the integration points the dynamic viscosity reads  $\mu^m = N_a^m \mu_a$ .

### 3.1.2 Solute transport equation

The mass fraction of the solute  $\omega$  is the primary unknown in the solute transport equation. The trial function of the mass fraction field  $\hat{\omega}$  approximates the mass fraction field as

$$\omega(\mathbf{x}, t) \approx \hat{\omega}(\mathbf{x}, t) = \omega_b(t) N_b(\mathbf{x}), \quad b = 1, \dots, n_n. \quad (3.29)$$

The finite element method discretizes the spatial domain and the nodal values of the mass fraction vary in time. The finite element basis functions  $N_b$  interpolate the nodal values over the flow domain. In this expression  $n_n$  denotes the number of finite element nodes.

#### Problem definition

The solute transport equation (2.73) written in index notation on a Cartesian field reads

$$nS\rho \frac{\partial \omega}{\partial t} = -\rho q_i \frac{\partial \omega}{\partial x_i} + \frac{\partial}{\partial x_i} \left( \rho D_{ij} \frac{\partial \omega}{\partial x_j} \right) + \tilde{\rho} (\tilde{\omega} - \omega) \bar{Q} \quad \text{on } \Omega, \quad (3.30)$$

where  $q_i$  denotes the components of the volumetric Darcy flux vector, and  $D_{ij}$  expresses the components of the hydrodynamic dispersion tensor. The volumetric source term  $\bar{Q} = \bar{Q}(\mathbf{x}, t)$  is positive for flow into the domain. A local positive value simulates an infiltration well and a negative value simulates an abstraction well. The injected solute mass is predefined as  $\tilde{\rho} \tilde{\omega} \bar{Q} = \bar{\rho} \bar{\omega} \bar{Q}$  for a positive value of the volumetric source (injection). For a negative value of the volumetric source the abstracted mass fraction follows from  $\tilde{\omega} = \omega$ , and the density of the fluid reads  $\tilde{\rho} = \rho$ . So, for negative values of the volumetric source the third term on the right hand side disappears.

The volumetric convective flux follows from Darcy's equation (2.26) as

$$q_i = -\frac{k_r K_{ij}}{\mu} \left( \frac{\partial p}{\partial x_j} - \rho g_j \right). \quad (3.31)$$

The coefficients in this equation were previously explained. The hydrodynamic dispersion tensor coefficients follow from equation (2.28), and reads

$$D_{ij} = nSD_c\delta_{ij} + (\alpha_l - \alpha_t)\frac{q_iq_j}{q} + \alpha_tq\delta_{ij}, \quad (3.32)$$

where  $q$  is the magnitude of the Darcy flux. The longitudinal dispersivity  $\alpha_l = \alpha_l(\mathbf{x})$  and transverse dispersivity  $\alpha_t = \alpha_t(\mathbf{x})$  are assumed to be time independent, and vary in space only. The molecular diffusivity  $D_c$  is assumed to be constant and includes the micro tortuosity of the soil structure.

The initial condition for the solute transport equation is given by a mass fraction field, expressed as

$$\omega = \omega^0 \quad \text{on } \Omega, \quad (3.33)$$

where  $\Omega$  denotes the flow domain,  $\Gamma$  will express the boundary of the domain.

First-type boundary conditions or Dirichlet boundary conditions specify the mass fraction on parts of the flow domain boundary as

$$\omega = \bar{\omega} \quad \text{on } \Gamma_1^\omega. \quad (3.34)$$

Robin or third-type boundary conditions express the mass flux into the flow domain as

$$\rho D_{ij} \frac{\partial \omega}{\partial x_j} n_i = \bar{\rho} \bar{q}_n (\bar{\omega} - \omega) \quad \text{on } \Gamma_2^\omega. \quad (3.35)$$

Here  $\bar{q}_n$  specifies an influx. The constraint poses a Robin condition for flow into the domain and reduces to a Von Neumann condition, which sets the dispersive flux to zero, for outflow.

### Spatial discretization

A weak formulation of the solute transport equation (3.30) over a single element reads

$$\begin{aligned} \int_{\Omega^e} W_a n S \rho \frac{\partial \hat{\omega}}{\partial t} d\Omega &= - \int_{\Omega^e} W_a \rho q_i \frac{\partial \hat{\omega}}{\partial x_i} d\Omega \\ &+ \int_{\Omega^e} W_a \frac{\partial}{\partial x_i} \left( \rho D_{ij} \frac{\partial \hat{\omega}}{\partial x_j} \right) d\Omega + \int_{\Omega^e} W_a \bar{\rho} (\bar{\omega} - \hat{\omega}) \bar{Q} d\Omega. \end{aligned} \quad (3.36)$$

The size of this local set of equations equals the number of nodes on the element. The weight function  $W_a$  is associated with node  $a$  of element  $e$ . Application of Green's theorem to the third term of equation (3.36) gives

$$\begin{aligned} \int_{\Omega^e} W_a n S \rho \frac{\partial \hat{\omega}}{\partial t} d\Omega &= - \int_{\Omega^e} W_a \rho q_i \frac{\partial \hat{\omega}}{\partial x_i} d\Omega - \int_{\Omega^e} \rho D_{ij} \frac{\partial W_a}{\partial x_i} \frac{\partial \hat{\omega}}{\partial x_j} d\Omega \\ &+ \int_{\Gamma^e} W_a \rho D_{ij} \frac{\partial \hat{\omega}}{\partial x_j} n_i d\Gamma + \int_{\Omega^e} W_a \bar{\rho} (\bar{\omega} - \hat{\omega}) \bar{Q} d\Omega. \end{aligned} \quad (3.37)$$

Here the order of integration has been reduced. Substitution of the interpolation function for the mass fraction trial function and replacing the weight functions by the basis functions according to Galerkin's principle gives

$$\begin{aligned} \int_{\Omega^e} N_a n S \rho \frac{d\hat{\omega}}{dt} d\Omega &= - \int_{\Omega^e} \rho q_i N_a \frac{\partial N_b}{\partial x_i} \omega_b d\Omega - \int_{\Omega^e} \rho D_{ij} \frac{\partial N_a}{\partial x_i} \frac{\partial N_b}{\partial x_j} \omega_b d\Omega \\ &+ \int_{\Gamma_2^e} N_a N_b (\tilde{\rho}(\tilde{\omega} - \omega) \bar{q}_n)_b d\Gamma + \int_{\Omega^e} N_a N_b (\tilde{\rho}(\tilde{\omega} - \omega) \bar{Q})_b d\Omega, \end{aligned} \quad (3.38)$$

where the weak boundary condition disappears for a prescribed outflow flux.

### Integration in time

The diffusive form of the solute transport equation gives rise to the following time dependent terms

$$\begin{aligned} \frac{d}{dt} (S\rho\hat{\omega}) &\approx \frac{(S\rho\hat{\omega}^{n+\theta})|^{p^{n+1}} - (S\rho\hat{\omega}^{n+\theta})|^{p^n}}{\hat{p}^{n+1} - \hat{p}^n} \frac{d\hat{p}}{dt} \\ &+ \frac{(S\rho\hat{\omega}^{n+1})|\omega^{n+1} - (S\rho\hat{\omega}^n)|\omega^n}{\hat{\omega}^{n+1} - \hat{\omega}^n} \frac{d\hat{\omega}}{dt} + \frac{(S\rho\hat{\omega}^{n+\theta})|^{T^{n+1}} - (S\rho\hat{\omega}^{n+\theta})|^{T^n}}{\hat{T}^{n+1} - \hat{T}^n} \frac{d\hat{T}}{dt}. \end{aligned} \quad (3.39)$$

Multiplication of the semi-discrete flow equation (3.10) by  $\hat{\omega}^{n+\theta}$ , and subtracting this expression from the diffusive form gives the convective formulation of the transport equation. The time dependent term for this formulation reads

$$\begin{aligned} (S\rho) \frac{d\hat{\omega}}{dt} &\approx \frac{(S\rho)|^{\omega^{n+1}} \hat{\omega}^{n+1} - (S\rho)|^{\omega^n} \hat{\omega}^n}{\hat{\omega}^{n+1} - \hat{\omega}^n} \frac{d\hat{\omega}}{dt} \\ &- \frac{(S\rho)|^{\omega^{n+1}} \hat{\omega}^{n+\theta} - (S\rho)|^{\omega^n} \hat{\omega}^{n+\theta}}{\hat{\omega}^{n+1} - \hat{\omega}^n} \frac{d\hat{\omega}}{dt}. \end{aligned} \quad (3.40)$$

According to this expression it follows that for  $\theta = 1$  the mass term should be evaluated as  $(S\rho)^n$ , and for  $\theta = 0$  the mass term should hold  $(S\rho)^{n+1}$ . For this reason, secants in the flow equation should be evaluated in the point  $(p^0, \omega^0, T^0)$  if the convective form is integrated implicitly, and in the point  $(p^1, \omega^1, T^1)$  if the convective form is integrated explicitly. It is noted that the secant procedure aims to find the change in  $(S\rho)$  over a time step for the flow equation. The convective form applies its actual value, which should be calculated at the end of the time step for an implicit formulation and at the start of the time step for an explicit form.

A global set of ordinary differential equations follows from the assemble of local sets of ordinary differential equations and can be written as

$$M_{ab} \frac{d\omega_b}{dt} = S_{ab} \omega_b + F_a. \quad (3.41)$$

The components of the capacity matrix  $M_{ab}$  [ml<sup>na-3</sup>] read

$$M_{ab} = \underset{e=1}{\overset{n_e}{\mathbf{A}}} \int_{\Omega^e} n S \rho N_a N_b d\Omega, \quad (3.42)$$

and its approximation

$$M_{ab} \approx \sum_{e=1}^{n_e} w_m^e (n S \rho N_a N_b |J|)_m^e. \quad (3.43)$$

This expression approximates the integrals by numerical integration. The capacity matrix has a consistent form. The components of the conductivity matrix  $S_{ab}$  [ $\text{ml}^{\text{nd}-3}\text{t}^{-1}$ ] are given by

$$\begin{aligned} S_{ab} = & - \sum_{e=1}^{n_e} \int_{\Omega^e} \rho q_i N_a \frac{\partial N_b}{\partial x_i} d\Omega - \sum_{e=1}^{n_e} \int_{\Omega^e} \rho D_{ij} \frac{\partial N_a}{\partial x_i} \frac{\partial N_b}{\partial x_j} d\Omega \\ & - \sum_{e=1}^{n_e} \int_{\Gamma_2^e} N_a N_b (\overline{\rho q_n})_b d\Gamma - \sum_{e=1}^{n_e} \int_{\Omega^e} N_a N_b (\overline{\rho Q})_b d\Omega. \end{aligned} \quad (3.44)$$

Numerical integration and local-global mapping obtains

$$\begin{aligned} S_{ab} \approx & - \sum_{e=1}^{n_e} w_m^e \left( \rho q_i N_a \frac{\partial N_b}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} |J| \right)_m^e - \sum_{e=1}^{n_e} w_m^e \left( \rho D_{ij} \frac{\partial N_a}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_b}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} |J| \right)_m^e \\ & - \sum_{e=1}^{n_{b2}} w_m^e (N_a N_b |J|)_m^e (\overline{\rho q_n})_b - (\overline{\rho Q})_a. \end{aligned} \quad (3.45)$$

Here  $n_{b2}$  denotes the number of nodes on an inflow boundary. The inverse of the Jacobian matrix multiplies the derivatives of the basis functions on the mapped local element in order to obtain the derivatives on the global element. The determinant of the Jacobian matrix  $|J|$  maps the length, surface area or volume of the local element to the global element. The prescribed volumetric flux  $N_b \overline{Q}_b$  is replaced by a point source  $\overline{Q}_a$ , which simulates wells. The force vector components  $F_a$  [ $\text{ml}^{\text{nd}-3}\text{t}^{-1}$ ] follow from

$$F_a = \sum_{e=1}^{n_e} \int_{\Gamma_2^e} N_a N_b (\overline{\rho \omega q_n})_b d\Gamma + \sum_{e=1}^{n_e} \int_{\Omega^e} N_a N_b (\overline{\rho \omega Q})_b d\Omega, \quad (3.46)$$

which is approximated by

$$F_a \approx \sum_{e=1}^{n_{b2}} w_m^e (N_a N_b |J|)_m^e (\overline{\rho \omega q_n})_b + (\overline{\rho \omega Q})_a. \quad (3.47)$$

The boundary flux term is only defined for boundary segments where  $\overline{q_n} > 0$  holds.

Linear  $\theta$ -weighted time integration transfers the set of ordinary differential equations to a set of algebraic equations. This procedure modifies equation (3.41) to

$$M_{ab}^{n+\theta} \frac{\omega_b^{n+1} - \omega_b^n}{\Delta t} = \theta S_{ab}^{n+1} \omega_b^{n+1} + (1 - \theta) S_{ab}^n \omega_b^n + \theta F_a^{n+1} + (1 - \theta) F_a^n. \quad (3.48)$$

Implicit formulations need updated computation of the capacity matrix components  $M_{ab}^{n+\theta}$ , the conductivity matrix components  $S_{ab}^{n+1}$  and the force vector entries  $F_a^{n+1}$ . The components follow from the yet unknown values of the primary variables. A Picard iteration scheme resolves the nonlinearities by

$$\begin{aligned} & \left( M_{ab}^{n+\theta} - \theta \Delta t S_{ab}^{n+1,r} \right) \omega_b^{n+1,r+1} \\ & = \left( M_{ab}^{n+\theta} + (1 - \theta) \Delta t S_{ab}^{n,r} \right) \omega_b^n + \theta \Delta t F_a^{n+1,r} + (1 - \theta) \Delta t F_a^n. \end{aligned} \quad (3.49)$$

Here  $r$  denotes the current nonlinear iteration and  $r + 1$  is the new iteration step. At each step a linearized set of equations needs to be solved.

### Algebraic equations

For an explicit integration scheme ( $\theta = 0$ ) the nonlinearities in the set of equations disappear. For pure convective flow the Galerkin explicit time integration schemes are unconditionally unstable [128]. For this reason, an implicit time integration scheme is preferred. The set of algebraic equations reads

$$\left(M_{ab}^{n+1,r} - \Delta t S_{ab}^{n+1,r}\right) \omega_b^{n+1} = M_{ab}^n \omega_b^n + \Delta t F_a^{n+1,r}. \quad (3.50)$$

This set of equations can be rewritten as

$$A_{ab}^\omega \omega_b = B_a^\omega, \quad (3.51)$$

where  $\omega_b$  is short for  $\omega_b^{n+1}$ . The stiffness matrix  $A_{ab}^\omega$  [ml<sup>nd-3</sup>] reads

$$\begin{aligned} A_{ab}^\omega = & \sum_{e=1}^{n_e} w_m^e (N_a N_b N_c |J|)_m^e (n S \rho)_c^{n+1} + \Delta t \sum_{e=1}^{n_e} w_m^e \left( (\rho q_i)^{n+1} N_a \frac{\partial N_b}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} |J| \right)_m^e \\ & + \Delta t \sum_{e=1}^{n_e} w_m^e \left( (\rho D_{ij})^{n+1} \frac{\partial N_a}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_b}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} |J| \right)_m^e \\ & + \Delta t \sum_{e=1}^{n_{b2}} w_m^e (N_a N_b |J|)_m^e (\overline{\rho q_n})_b^{n+1} + \Delta t (\overline{\rho Q_\delta})_a^{n+1}, \quad (3.52) \end{aligned}$$

and the force vector  $B_a^\omega$  [ml<sup>nd-3</sup>] is given by

$$\begin{aligned} B_a^\omega = & \sum_{e=1}^{n_e} w_m^e (N_a N_b N_c |J|)_m^e \omega_b^n (n S \rho)_c^{n+1} \\ & + \Delta t \sum_{e=1}^{n_{b2}} w_m^e (N_a N_b |J|)_m^e (\overline{\rho q_n \bar{\omega}})_b^{n+1} + \Delta t (\overline{\rho Q_\delta \bar{\omega}})_a^{n+1}. \quad (3.53) \end{aligned}$$

Here, second-type boundaries only contribute, if the flux prescribes flow into the flow domain.

The Darcy (mass) flux components in an integration point can be found by

$$(\rho q_i)_m = - \left[ \frac{\rho k_r K_{ij}}{\mu} \left( \frac{\partial N_a}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_j} p_a - N_a (\rho g_j^x)_a \right) \right]_m. \quad (3.54)$$

This expression interpolates the pressure with a lower order interpolation function than the density scaled gravitational vector. As a result, the velocity distribution will not be consistent over the element. For the convective term the velocity will be integrated over the element and the error in the approximation will cancel out. But this is not the case for the dispersive term and Voss and Souza [118] proposed a second order reduction scheme for the interpolation of the gravity term. They first wrote the gravity vector in local coordinates for node  $a$  as

$$(\rho g_i^\xi)_a = \left( \frac{\partial x_j}{\partial \xi_i} \rho g_j^x \right)_a. \quad (3.55)$$

Secondly, they interpolated the nodal vectors over the interior of the elements by the scaled derivative of the interpolation function. This interpolation reads

$$\rho g_i^\xi = N_a^i \left( \rho g_i^\xi \right)_a, \quad (3.56)$$

where no summation over  $i$  takes place. The modified interpolation function in the  $i^{\text{th}}$  direction are found, according to

$$N_a^i = r^i \left| \frac{\partial N_a}{\partial \xi_i} \right|, \quad r^i = 1 / \sum_a \left| \frac{\partial N_a}{\partial \xi_i} \right|, \quad (3.57)$$

where  $r^i$  forces the sum of the reduced interpolation functions to be one, on each location. The global gravity force vector components follow from the local components by multiplying them by the Jacobian matrix. This procedure reads

$$\rho g_i^x = \frac{\partial \xi_j}{\partial x_i} \rho g_j^\xi. \quad (3.58)$$

A consistent velocity field finally follows from

$$\begin{aligned} (\rho q_i)_m = & - \left( K_{ij} N_b \frac{\partial \xi_k}{\partial x_j} \frac{\partial N_a}{\partial \xi_k} \right)_m p_a \left( \frac{\rho k_r}{\mu} \right)_b \\ & + \left( K_{ij} N_b \frac{\partial \xi_k}{\partial x_j} N_a^k \right)_m \left( \rho g_k^\xi \right)_a \left( \frac{\rho k_r}{\mu} \right)_b, \end{aligned} \quad (3.59)$$

where no summation over  $k$  in the second term takes place.

The mass equivalent of the hydrodynamic dispersion  $D_{ij}$  in integration point  $m$  follows from

$$(\rho D_{ij})_m = (\rho n S)_m D_c \delta_{ij} + \left( (\alpha_l - \alpha_t) N_a \frac{q_i q_j}{q} \right)_m \rho_a + (\alpha_t q N_a)_m \rho_a \delta_{ij}, \quad (3.60)$$

where the molecular dispersion  $D_c$  is constant, the longitudinal dispersivity  $\alpha_l$  and transverse dispersivity  $\alpha_t$  are constant over the element. The components of the Darcy flux vector  $(q_i)_m$  and its magnitude  $(q)_m$  follow from equation (3.59).

### 3.1.3 Heat transport equation

Temperature  $T$  is the primary unknown in the heat transport equation. The finite element formulation approximates the temperature by a trial function  $\hat{T}$ . The nodal values of the trial functions  $T_b$  vary in time, finite element basis functions interpolate these nodal values in space, as

$$T(\mathbf{x}, t) \approx \hat{T}(\mathbf{x}, t) = T_b(t) N_b(\mathbf{x}). \quad b = 1, \dots, n_n, \quad (3.61)$$

Here  $N_b$  denotes the basis function or shape function associated with node  $b$ , and  $n_n$  is the number of nodes in the finite element mesh.

### Problem definition

The heat transport equation (2.74), reformulated in index notation on Cartesian coordinates reads

$$nS\rho c \frac{\partial T}{\partial t} + (1-n)\rho^s c^s \frac{\partial T}{\partial t} = -\rho c q_i \frac{\partial T}{\partial x_i} + \frac{\partial}{\partial x_i} \left( H_{ij} \frac{\partial T}{\partial x_j} \right) + c\tilde{\rho} (\tilde{T} - T) \bar{Q} \quad \text{on } \Omega, \quad (3.62)$$

where  $c$  denotes the specific heat capacity of the fluid phase,  $c^s = c^s(\mathbf{x})$  is the specific heat capacity of the solid phase, and  $H_{ij} = H_{ij}(\mathbf{x}, t)$  denotes the hydrodynamic thermal dispersion tensor components. If the volumetric source term specifies outflow then  $\tilde{T} = T$  and the source term disappears from the equation. Thermal energy still leaves the system at these points at a rate  $c\rho T Q$ .

The hydrodynamic thermal dispersion  $H_{ij}$  includes the mechanical part of the hydrodynamic dispersion tensor, and conductivity of the partly saturated soil. The coefficients of the hydrodynamic thermal dispersion tensor follow from equation (2.43), and are written as

$$H_{ij} = nS\lambda\delta_{ij} + (1-n)\lambda_{ij}^s + c\rho \left( (\alpha_l - \alpha_t) \frac{q_i q_j}{q} + \delta_{ij} \alpha_t q \right). \quad (3.63)$$

The thermal components of the soil skeleton  $\lambda_{ij}^s = \lambda_{ij}^s(\mathbf{x})$  vary in space, but are assumed to be constant in time. The water phase conductivity  $\lambda$  is assumed to be a constant scalar variable.

The initial condition is given by an initial temperature distribution in space, written as

$$T = T^0 \quad \text{on } \Omega. \quad (3.64)$$

A temperature distribution at equilibrium follows from a boundary value problem where the conditions on the boundary  $\Gamma$  of the flow domain  $\Omega$  need to be specified.

First-type boundary conditions or Dirichlet boundary conditions prescribe the temperature on parts of the flow boundary, according to

$$T = \bar{T} \quad \text{on } \Gamma_1^T. \quad (3.65)$$

Robin or third-type boundary conditions specify the derivative of the temperature over the boundary, which corresponds to an energy flux. This condition reads

$$H_{ij} \frac{\partial T}{\partial x_j} n_i = c\tilde{\rho} \bar{q}_n (\tilde{T} - T) \quad \text{on } \Gamma_2^T, \quad (3.66)$$

where the volumetric boundary flux  $\bar{q}_n$  is positive for inflow. For this condition the temperature of the fluid reads  $\tilde{T} = \bar{T}$ , and the fluid density is prescribed as  $\tilde{\rho} = \bar{\rho}$ . If the boundary flux has a negative value then the dispersive flux over the boundary is set to zero.

### Spatial discretization

Multiplication of the heat transport equation (3.62) by a weighting function and integration over a single element gives

$$\begin{aligned} \int_{\Omega^e} W_a n S \rho c \frac{\partial \hat{T}}{\partial t} d\Omega + \int_{\Omega^e} W_a (1-n) \rho^s c^s \frac{\partial \hat{T}}{\partial t} d\Omega = - \int_{\Omega^e} W_a \rho c q_i \frac{\partial \hat{T}}{\partial x_i} d\Omega \\ + \int_{\Omega^e} W_a \frac{\partial}{\partial x_i} \left( H_{ij} \frac{\partial \hat{T}}{\partial x_j} \right) d\Omega + \int_{\Omega^e} W_a c \tilde{\rho} (\tilde{T} - \hat{T}) \bar{Q} d\Omega, \end{aligned} \quad (3.67)$$

where  $a = 1, \dots, n_{ne}$  and  $n_{ne}$  denotes the number of nodal points in the element. This number equals the number of equations in the local set. Green's theorem on the fourth term of equation (3.67) reduces the order of integration. The result is written as

$$\begin{aligned} \int_{\Omega^e} W_a n S \rho c \frac{\partial \hat{T}}{\partial t} d\Omega + \int_{\Omega^e} W_a (1-n) \rho^s c^s \frac{\partial \hat{T}}{\partial t} d\Omega = - \int_{\Omega^e} W_a \rho c q_i \frac{\partial \hat{T}}{\partial x_i} d\Omega \\ - \int_{\Omega^e} H_{ij} \frac{\partial W_a}{\partial x_i} \frac{\partial \hat{T}}{\partial x_j} d\Omega + \int_{\Gamma^e} W_a H_{ij} \frac{\partial \hat{T}}{\partial x_j} n_i d\Gamma + \int_{\Omega^e} W_a c \tilde{\rho} (\tilde{T} - \hat{T}) \bar{Q} d\Omega. \end{aligned} \quad (3.68)$$

Substitution of the interpolation functions for the temperature trial function and replacing the weighting functions by the basis functions according to Galerkin's method, gives the local set of equations, which can be expressed as

$$\begin{aligned} \int_{\Omega^e} N_a n S \rho c \frac{d\hat{T}}{dt} d\Omega + \int_{\Omega^e} N_a (1-n) \rho^s c^s \frac{d\hat{T}}{dt} d\Omega \\ = - \int_{\Omega^e} N_a \rho c q_i \frac{\partial N_b}{\partial x_i} T_b d\Omega - \int_{\Omega^e} H_{ij} \frac{\partial N_a}{\partial x_i} \frac{\partial N_b}{\partial x_j} T_b d\Omega \\ + \int_{\Gamma^e} N_a N_b [c \tilde{\rho} (\tilde{T} - T) \bar{q}_n]_b d\Gamma + \int_{\Omega^e} N_a N_b [c \tilde{\rho} (\tilde{T} - T) \bar{Q}]_b d\Omega. \end{aligned} \quad (3.69)$$

In this expression, the boundary integral was replaced by the weak definition. Essential (first type) boundary conditions are included by a modification of the weighting functions.

### Integration in time

The global set of ordinary differential equations is written as

$$M_{ab} \frac{dT_b}{dt} = S_{ab} T_b + F_a. \quad (3.70)$$

The components of the consistent form of the capacity matrix  $M_{ab}$  [ $\text{ml}^{n_a-1} \text{t}^{-2} \text{T}^{-1}$ ] are given by

$$M_{ab} = \sum_{e=1}^{n_e} \int_{\Omega^e} N_a N_b (n S \rho c + (1-n) \rho^s c^s) d\Omega, \quad (3.71)$$

Numerical integration over local elements gives

$$M_{ab} \approx \sum_{e=1}^{n_e} w_m^e [N_a N_b (n S \rho c + (1-n) \rho^s c^s) |J|]_m^e. \quad (3.72)$$

Here numerical integration approximates the integral over the element. The local matrices are assembled in the global capacity matrix. The components of the conductivity matrix  $S_{ab}$  [ $\text{ml}^{n_d-1} \text{t}^{-3} \text{T}^{-1}$ ] follow from

$$S_{ab} = - \sum_{e=1}^{n_e} \int_{\Omega^e} \rho c q_i N_a \frac{\partial N_b}{\partial x_i} d\Omega - \sum_{e=1}^{n_e} \int_{\Omega^e} H_{ij} \frac{\partial N_a}{\partial x_i} \frac{\partial N_b}{\partial x_j} d\Omega \\ - \sum_{e=1}^{n_e} \int_{\Gamma_2^e} N_a N_b (c \bar{\rho} \bar{q}_n)_b d\Gamma - \sum_{e=1}^{n_e} \int_{\Omega^e} N_a N_b (c \bar{\rho} \bar{Q})_b d\Omega, \quad (3.73)$$

which is approximated by

$$S_{ab} \approx - \sum_{e=1}^{n_e} w_m^e \left( c \rho q_i N_a \frac{\partial N_b}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} |J| \right)_m^e - \sum_{e=1}^{n_e} w_m^e \left( H_{ij} \frac{\partial N_a}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_b}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} |J| \right)_m^e \\ - \sum_{e=1}^{n_{b2}} w_m^e c (N_a N_b |J|)_m^e (\bar{\rho} \bar{q}_n)_b - (c \bar{\rho} \bar{Q}_\delta)_a. \quad (3.74)$$

Here  $n_{b2}$  denotes the number of boundary nodes on the second-type inflow boundary. If outflow is prescribed then the term is excluded from the formulation. The volumetric source is replaced by a point source  $\bar{Q}_a$  located in node  $a$ , which is included only if the source has a positive value and energy is flowing into the domain. The components of the force vector  $F_a$  [ $\text{ml}^{n_d-1} \text{t}^{-3}$ ] read

$$F_a = \sum_{e=1}^{n_e} \int_{\Gamma_2^e} N_a N_b (c \bar{\rho} \bar{q}_n \bar{T})_b d\Gamma + \sum_{e=1}^{n_e} \int_{\Omega^e} N_a N_b (c \bar{\rho} \bar{Q} \bar{T})_b d\Omega, \quad (3.75)$$

or

$$F_a \approx \sum_{e=1}^{n_{b2}} w_m^e c (N_a N_b |J|)_m^e (\bar{\rho} \bar{q}_n \bar{T})_b + (c \bar{\rho} \bar{Q}_\delta \bar{T})_a. \quad (3.76)$$

Both terms only apply for inflow conditions ( $\bar{q}_n > 0$  and  $\bar{Q} > 0$ ); for outflow conditions the terms become zero.

Linear  $\theta$ -weighted time integration translates equation (3.70) into

$$M_{ab}^{n+\theta} \frac{T_b^{n+1} - T_b^n}{\Delta t} = \theta S_{ab}^{n+1} T_b^{n+1} + (1-\theta) S_{ab}^n T_b^n + \theta F_a^{n+1} + (1-\theta) F_a^n, \quad (3.77)$$

where  $0 < \theta < 1$ . In general, components of the capacity matrix  $M_{ab}^{n+\theta}$ , the conductivity matrix  $S_{ab}^{n+1}$ , and the force vector  $F_a^{n+1}$  follow from the yet unknown values of the primary variables. Successive substitution or Picard iteration resolves the nonlinearities by

$$\left( M_{ab}^{n+\theta} - \theta \Delta t S_{ab}^{n+1,r} \right) T_b^{n+1,r+1} \\ = \left( M_{ab}^{n+\theta} + (1-\theta) \Delta t S_{ab}^{n,r} \right) T_b^n + \theta \Delta t F_a^{n+1,r} + (1-\theta) \Delta t F_a^n, \quad (3.78)$$

where  $r$  denotes the current iteration step and  $r+1$  is the new step.

### Algebraic equations

The implicit formulation for the heat transport equation reads

$$\left(M_{ab}^{n+1,r} - \Delta t S_{ab}^{n+1,r}\right) T_b^{n+1} = M_{ab}^{n+1,r} T_b^n + \Delta t F_a^{n+1,r}. \quad (3.79)$$

The corresponding set of linear algebraic equation reads

$$A_{ab}^T T_b = B_a^T, \quad (3.80)$$

where  $T_b^{n+1}$  is short for  $T_b$ . The stiffness matrix components  $A_{ab}^T$  [ $\text{ml}^{n_d-1} \text{t}^{-2} \text{T}^{-1}$ ] read

$$\begin{aligned} A_{ab}^T = & \sum_{e=1}^{n_e} w_m^e (N_a N_b N_c |J|)_m^e (nS\rho c + (1-n)\rho^s c^s)_c^{n+1} \\ & + \Delta t \sum_{e=1}^{n_e} w_m^e c \left( (\rho q_i)^{n+1} N_a \frac{\partial N_b}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} |J| \right)_m^e + \Delta t \sum_{e=1}^{n_e} w_m^e \left( H_{ij}^{n+1} \frac{\partial N_a}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_b}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} |J| \right)_m^e \\ & + \Delta t \sum_{e=1}^{n_{b2}} w_m^e c (N_a N_b |J|)_m^e (\bar{\rho} q_n)_b^{n+1} + c \Delta t (\bar{\rho} \bar{Q}_\delta)_a^{n+1}, \quad (3.81) \end{aligned}$$

and the components of the right hand side force vector  $B_a^T$  [ $\text{ml}^{n_d-1} \text{t}^{-2}$ ] read

$$\begin{aligned} B_a^T = & \sum_{e=1}^{n_e} w_m^e (N_a N_b N_c |J|)_m^e T_b^n (nS\rho c + (1-n)\rho^s c^s)_c^{n+1} \\ & + \Delta t \sum_{e=1}^{n_{b2}} w_m^e c (N_a N_b |J|)_m^e (\bar{\rho} q_n \bar{T})_b^{n+1} + c \Delta t (\bar{\rho} \bar{Q}_\delta \bar{T})_a^{n+1}. \quad (3.82) \end{aligned}$$

The solution of this set of equation is found sequentially.

In order to set up the set of algebraic equations, the components of hydrodynamic thermal dispersion tensor  $H_{ij}$  need to be calculated in the integration points  $m$ . The components of the tensor follow from

$$\begin{aligned} (H_{ij})_m = & (nS)_m \lambda \delta_{ij} + ((1-n)\lambda_{ij}^s)_m \\ & + c \left( N_a (\alpha_l - \alpha_t) \frac{q_i q_j}{q} + \delta_{ij} N_a \alpha_t q \right)_m \rho_a, \quad (3.83) \end{aligned}$$

where the liquid phase specific heat capacity  $c$  and the conductivity  $\lambda$  are constant. The thermal conductivity coefficients for the soil skeleton  $\lambda_{ij}^s$  are constant in time and constant over the element.

## 3.2 Problem definition

The finite element method transforms the set of partial differential equations into a set of algebraic equations. The number of algebraic equations equals the number of nodes in the mesh times the degrees of freedom per node. Here three degrees of freedom for each node apply: pressure, mass fraction and temperature. The set of equations can

either be solved in its coupled full-size form or in a decomposed form by decoupling the equations per degree of freedom.

Section 3.2.1 presents the system of algebraic equations and the solution strategy for the partly linearized set of coupled algebraic equations. Section 3.2.2 proposes a procedure to include initial conditions and section 3.2.3 formulates set of physically admissible boundary conditions that combine the conditions given for each equation separately.

### 3.2.1 System of algebraic equations

The coupled solution procedure requires the formulation of the constitutive relation for fluid density and fluid viscosity as a function of all primary unknowns. Introduced into the set of equations this formulation produces coupling terms in the global conductivity matrix. The capacity matrix in the flow equation already contained coupling terms through the functional dependency of relations on mass fraction and temperature. The decoupled solution procedure accounts for the coupling terms by the force vector. Generally, the Darcy velocity and dispersivity are included in the set of transport equations. For this reason both coupled and decoupled equations have to be solved iteratively. In this thesis, the decoupled procedure is preferred as it saves computer memory and is able to deal with more complex constitutive relations.

At each time step, a linearized symmetric flow equation and two non-symmetric transport equations need to be solved, because an implicit time integration procedure is used for all equations. Coupling terms and nonlinearities are resolved sequentially. The amount of work depends on the nonlinearities present in each of the equations, on the coupling terms that have to be accounted for, and on the required time step size relative to the simulated process.

Stability requirements may reduce the allowable time step size. The time step size is also restricted to the range of the nonlinear solver and the mesh refinement strategy. The stability analysis of the set of coupled nonlinear equations is not easily performed, and only strongly simplified stability checks can be carried out. In this thesis, an implicit time integration formulation is proposed, which is less restrictive on the time step size. A Picard iteration scheme is preferred over a Newton-Raphson scheme, because the range of convergence of the first scheme is wider [100].

The proposed (decoupled) sequential procedure is expressed by

$$\begin{aligned} A_{ab}^p(p^n, p^r, \omega^r, T^r) p_b^{r+1} &= B_a^p(p^n, p^r, \omega^n, \omega^r, T^n, T^r), \\ A_{ab}^\omega(p^{r+1}, \omega^r, T^r) \omega_b^{r+1} &= B_a^\omega(p^{r+1}, \omega^n, \omega^r, T^r), \\ A_{ab}^T(p^{r+1}, \omega^{r+1}, T^r) T_b^{r+1} &= B_a^T(p^{r+1}, \omega^{r+1}, T^n, T^r). \end{aligned} \quad (3.84)$$

Here  $n$  indicates the current time step,  $r$  is short for the current iteration step at the new time step  $n + 1$  and  $r + 1$  denotes the new iteration step at new time step  $n + 1$ . The sequential procedure first solves the pressure equation, posed by the first line in equation (3.84). This line presents the set of algebraic equations that poses the flow problem. Both the left hand side matrix and the right hand side vector are based on primary unknowns evaluated on the current and the new time step. At

the new time step these unknowns are approximated by their values at the current nonlinear approximation at this time step or their values at the previous time step at the start of the nonlinear iteration. The expression for the system matrix coefficients was given by equation (3.22). The right hand side vector coefficients were expressed by equation (3.23). The Darcy velocity field follows from the updated pressure field, and will be imposed into the transport equations. The second line in equation (3.84) states the solute transport equation. System matrix coefficients and right hand side vector coefficients followed from equations (3.52) and (3.53). The updated mass fraction field is used in the heat transport computation. The third line in equation (3.84) gives the heat transport equation, for which the coefficients of the system matrix were given by equation (3.81) and the right hand side vector coefficients followed from equation (3.82). The sequential procedure is repeated until the solution of pressure, mass fraction and temperature stabilize or successive changes in the nodal values are within a prescribed tolerance. The algorithm switches to the next time level after it achieves convergence.

Implicit formulations of linear convection-diffusion problems are found to be unconditionally stable. This does not hold for the nonlinear flow problem. For nonlinear problems the time step size is restricted by the range of convergence of the nonlinear solver. The number of Picard iterations per time step provides a heuristic basis for an automatic time step algorithm. The proposed time step algorithm considers four situations. If the number of iterations needed to achieve convergence in a time step is less than a lower limit value ( $n_r < 5$ ), then the size of the next time step may be doubled according to  $\Delta t^{n+1} = 2.0 \Delta t^n$ . Dissipation of an infiltration front may involve this situation. If on the other hand the number of nonlinear iterations exceeds an upper limit value, ( $n_r > 10$ ), then the size of the next time step needs to be reduced by  $\Delta t^{n+1} = 0.5 \Delta t^n$ . If the number of nonlinear iterations exceeds a second upper limit value, ( $n_r > 20$ ), then the size of the current time step will be reduced once more according to  $\Delta t^n = 0.5 \Delta t^n$ . This situation may occur if boundary conditions change or infiltration fronts propagate into a new material while the time step is relatively large.

### Stability and accuracy

This section considers the stability and accuracy of a one-dimensional convection-diffusion test problem, written as

$$C \frac{\partial u}{\partial t} + U \frac{\partial u}{\partial x} - D \frac{\partial^2 u}{\partial x^2} = 0, \quad \text{on } \Omega. \quad (3.85)$$

This expression does not include source terms and simplifies the previously discussed flow and transport equations. The flow equation (3.2) reduces to this form if fluid density, fluid viscosity, intrinsic permeability of the soil structure and saturation and relative permeability of the subsoil do not vary in time and space. For this case the coefficients read  $C = \alpha_p \rho S$  and  $D = \rho k_r K_{ij} / \mu$ . The solute transport equation (3.30) complies to the test problem if Fick's law for diffusion holds and density changes are not taken into account. The convection-diffusion coefficients then read  $C = nS\rho$ ,  $U = \rho q_i$  and  $D = \rho D_{ij}$ . The test problem simplifies the heat transport equation (3.62) for  $C = (nS\rho c + (1-n)\rho^s c^s)$ ,  $U = \rho c q_i$  and  $D = \rho H_{ij}$ . The thermal conductivity of

the fluid phase and the solid phase both contribute to the constant diffusion coefficient. Hydrodynamic dispersion is disregarded for the stability and accuracy analysis. The unknown  $u$  in the test problem, either specifies pressure  $p$ , mass fraction  $\omega$  or temperature  $T$ .

Spatial discretization and integration in time lead to algebraic equations (3.18), (3.48) and (3.77). These algebraic equations simplify the test problem to

$$M_{ab}^{n+\theta} \frac{u_b^{n+1} - u_b^n}{\Delta t} = \theta S_{ab}^{n+1} u_b^{n+1} + (1 - \theta) S_{ab}^n u_b^n. \quad (3.86)$$

In this section the convective term and the diffusive term will be considered separately. The consistent-mass capacity matrix according to equations (3.13), (3.43) and (3.72) reads

$$M_{ab} = \mathbf{A}_{e=1}^{n_e} \int_{\Omega^e} C N_a N_b d\Omega. \quad (3.87)$$

The diffusive part of the conductivity matrix  $S_{ab}^D$  according to equation (3.15), (3.45) and (3.74) reads

$$S_{ab}^D = - \mathbf{A}_{e=1}^{n_e} \int_{\Omega^e} D \frac{\partial N_a}{\partial x} \frac{\partial N_b}{\partial x} d\Omega. \quad (3.88)$$

The convective part of the conductivity matrix  $S_{ab}^U$  follows from the numerical derivation of transport equations (3.45) and (3.74) as

$$S_{ab}^U = - \mathbf{A}_{e=1}^{n_e} \int_{\Omega^e} U N_a \frac{\partial N_b}{\partial x} d\Omega. \quad (3.89)$$

The nodal equation for the test problem on a uniform grid reads

$$\begin{aligned} & \frac{\Delta x C}{6} (u_{j-1}^{n+1} + 4u_j^{n+1} + u_{j+1}^{n+1}) - \frac{\Delta x C}{6} (u_{j-1}^n + 4u_j^n + u_{j+1}^n) \\ & + \theta \frac{\Delta t U}{2} (-u_{j-1}^{n+1} + u_{j+1}^{n+1}) + (1 - \theta) \frac{\Delta t U}{2} (-u_{j-1}^n + u_{j+1}^n) \\ & + \theta \frac{D \Delta t}{\Delta x} (-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}) + (1 - \theta) \frac{D \Delta t}{\Delta x} (-u_{j-1}^n + 2u_j^n - u_{j+1}^n) = 0. \end{aligned} \quad (3.90)$$

The finite difference form follows from this equation by mass lumping and multiplication by  $1/\Delta x \Delta t$ . This form reads

$$\begin{aligned} & \frac{C}{\Delta t} u_j^{n+1} - \frac{C}{\Delta t} u_j^n + \theta \frac{U}{2\Delta x} (-u_{j-1}^{n+1} + u_{j+1}^{n+1}) + (1 - \theta) \frac{U}{2\Delta x} (-u_{j-1}^n + u_{j+1}^n) \\ & + \theta \frac{D}{\Delta x^2} (-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}) + (1 - \theta) \frac{D}{\Delta x^2} (-u_{j-1}^n + 2u_j^n - u_{j+1}^n) = 0. \end{aligned} \quad (3.91)$$

Both finite element and finite difference formulation solve a modified equation, which follows from a Taylor series expansion with reference point at  $(j\Delta x, n\theta\Delta t)$ . The expansion of the discrete unknown  $u_j^{n+1}$  in time gives

$$u_j^{n+1} \approx u_j^{n+\theta} + (1 - \theta)\Delta t \left. \frac{\partial u}{\partial t} \right|_{u_j^{n+\theta}} + (1 - \theta)^2 \frac{\Delta t^2}{2} \left. \frac{\partial^2 u}{\partial t^2} \right|_{u_j^{n+\theta}} + O(\Delta t^3). \quad (3.92)$$

The expansion of  $u_{j+1}^n$  in space and time reads

$$\begin{aligned} u_{j+1}^n \approx & u_j^{n+\theta} + \Delta x \left. \frac{\partial u}{\partial x} \right|_{u_j^{n+\theta}} - \theta \Delta t \left. \frac{\partial u}{\partial t} \right|_{u_j^{n+\theta}} + \frac{\Delta x^2}{2} \left. \frac{\partial^2 u}{\partial x^2} \right|_{u_j^{n+\theta}} \\ & + \theta \Delta x \Delta t \left. \frac{\partial^2 u}{\partial x \partial t} \right|_{u_j^{n+\theta}} + \theta^2 \frac{\Delta t^2}{2} \left. \frac{\partial^2 u}{\partial t^2} \right|_{u_j^{n+\theta}} + O(\Delta x^3) + O(\Delta t^3). \end{aligned} \quad (3.93)$$

The modified equation for the consistent-mass discretization with reference point at  $(j\Delta x, n\theta\Delta t)$  reads

$$\begin{aligned} C \frac{\partial u}{\partial t} + U \frac{\partial u}{\partial x} - D \frac{\partial^2 u}{\partial x^2} = & -(1-2\theta) \frac{C\Delta t}{2} \frac{\partial^2 u}{\partial t^2} - \frac{C\Delta x^2}{6} \frac{\partial^3 u}{\partial t \partial x^2} \\ & - \frac{C\Delta t^2}{6} \frac{\partial^3 u}{\partial t^3} - \frac{U\Delta x^2}{6} \frac{\partial^3 u}{\partial x^3} + \frac{D\Delta x^2}{12} \frac{\partial^4 u}{\partial x^4} + \theta(1-\theta) \frac{C\Delta t^2}{2} \frac{\partial^3 u}{\partial t^3} \\ & - \theta(1-\theta) \frac{U\Delta t^2}{2} \frac{\partial^3 u}{\partial t^2 \partial x} + \theta(1-\theta) \frac{D\Delta t^2}{3} \frac{\partial^4 u}{\partial t^2 \partial x^2} + \dots \end{aligned} \quad (3.94)$$

The right hand side presents the truncation errors, which are in the order of  $O(\Delta t, \Delta x^2)$  for the fully-implicit and explicit form, and correspond to  $O(\Delta t^2, \Delta x^2)$  for  $\theta = \frac{1}{2}$ . The first order derivative term in time introduces numerical diffusion [119]. The time derivative converts into spatial derivatives by

$$C \frac{\partial^2 u}{\partial t^2} = -\frac{\partial}{\partial t} \left( U \frac{\partial u}{\partial x} - D \frac{\partial^2 u}{\partial x^2} \right) = \frac{U^2}{C} \frac{\partial^2 u}{\partial x^2} - \frac{2UD}{C} \frac{\partial^3 u}{\partial x^3} + \frac{D^2}{C} \frac{\partial^4 u}{\partial x^4}. \quad (3.95)$$

The numerical diffusion, introduced by integration in time, reads

$$D_{num} = \left(\theta - \frac{1}{2}\right) \frac{U^2 \Delta t}{C}. \quad (3.96)$$

This artificial diffusion contributes to the diffusion  $D$  in equation (3.85). The actual velocity and actual diffusion in this equation hold  $U/C$  and  $D/C$ . For an explicit scheme the numerical diffusion equals  $D_{num} = -\frac{1}{2}U^2\Delta t/C$ , whereas for a fully implicit scheme this term reads  $D_{num} = \frac{1}{2}U^2\Delta t/C$ . For stability the total diffusion should be positive,  $D + D_{num} > 0$ . A positive value of the numerical diffusion stabilizes the method and adds diffusion to the system. The stability requirement poses a restriction on the time step for the explicit formulation according to

$$\Delta t < 2CD/U^2. \quad (3.97)$$

For a time weighting factor  $\frac{1}{2} \leq \theta \leq 1$  the total diffusion is positive independent of the time step. For accuracy  $D \gg D_{num}$  in order to obtain accurate solutions. The approximate equation for the lumped-mass form gives

$$C \frac{\partial u}{\partial t} + U \frac{\partial u}{\partial x} - D \frac{\partial^2 u}{\partial x^2} = -(1-2\theta) \frac{C\Delta t}{2} \frac{\partial^2 u}{\partial t^2} - \frac{U\Delta x^2}{6} \frac{\partial^3 u}{\partial x^3} + \frac{D\Delta x^2}{12} \frac{\partial^4 u}{\partial x^4} + \dots \quad (3.98)$$

For  $\theta = 0$  this form corresponds to the forward in time central in space finite difference approximation, a backward in time central in space finite difference formulation follows from  $\theta = 1$ .

Alternatively Von Neumann analysis considers the stability in terms of Fourier series. The initial condition has to be developed into a Fourier series of cosine functions with various wave lengths  $L$ . The initial condition reads  $u(x, 0) = u_0 \cos(kx)$ , where  $k = 2\pi/L$  will be represented by the real part of

$$u(x, 0) = u_0 \exp(ikx). \quad (3.99)$$

In a grid point the initial value reads

$$u_j = u_0 \exp(ij\xi), \quad (3.100)$$

where  $\xi = k\Delta x = 2\pi\Delta x/L$ . The smallest grid size is  $\Delta x = 0$  and the largest grid size is  $\Delta x = L/2$  as a wave can only be represented by at least two grid points. For this reason  $0 \leq \xi \leq \pi$  holds. In the adjacent grid point the discrete value of the unknown equals

$$u_{j+1}^n = u_0 \exp(i\xi) \exp(ij\xi). \quad (3.101)$$

For the next time step the grid point value reads

$$u_j^{n+1} = u_0 \rho \exp(ij\xi), \quad (3.102)$$

where  $\rho$  denotes the amplification factor. The modulus of the amplification factor  $|\rho|$  multiplies the amplitude of the initial signal. The requirement for stability therefore states  $|\rho| \leq 1$ . The phase shift follows from the argument of the amplification factor, which will be written as  $\arg(\rho)$ . The relative celerity ratio is the ratio of arguments of the amplification factor at different time steps and expresses the ratio of phase angles [128] by  $\arg(\rho)L/2\pi U\Delta t$ . Both amplitude ratio and relative celerity ratio should be close to unity for accuracy. The discrete grid point expressions entered in the nodal consistent-mass equation (3.90) give the expression for the amplification factor for the test problem. Using  $\{\exp(i\xi) + \exp(-i\xi)\}/2 = \cos \xi$  and  $\{\exp(i\xi) - \exp(-i\xi)\}/2i = \sin \xi$  the expression for the amplification factor reads

$$\rho = \frac{(\cos \xi + 2) + 3(1 - \theta)\lambda(\cos \xi - 1) - 3\sigma i(1 - \theta)\sin \xi}{(\cos \xi + 2) - 3\theta\lambda(\cos \xi - 1) + 3\sigma i\theta\sin \xi}, \quad (3.103)$$

where  $\lambda = 2D\Delta t/C\Delta x^2$ , and  $\sigma = U\Delta t/C\Delta x$ . Figure 3.1 show the amplitude ratio and relative celerity ratio for a number of combinations of  $\lambda$  and  $\sigma$ .

The amplification factor for the lumped-mass formulation, given by equation (3.91), reads

$$\rho = \frac{1 + (1 - \theta)\lambda(\cos \xi - 1) - \sigma(1 - \theta)i\sin \xi}{1 - \theta\lambda(\cos \xi - 1) + \sigma\theta i\sin \xi}. \quad (3.104)$$

According to these expressions both methods are unconditionally stable for  $\frac{1}{2} \leq \theta \leq 1$ . The explicit mass-lumped method is stable for

$$\sigma^2 \leq \lambda \leq 1. \quad (3.105)$$

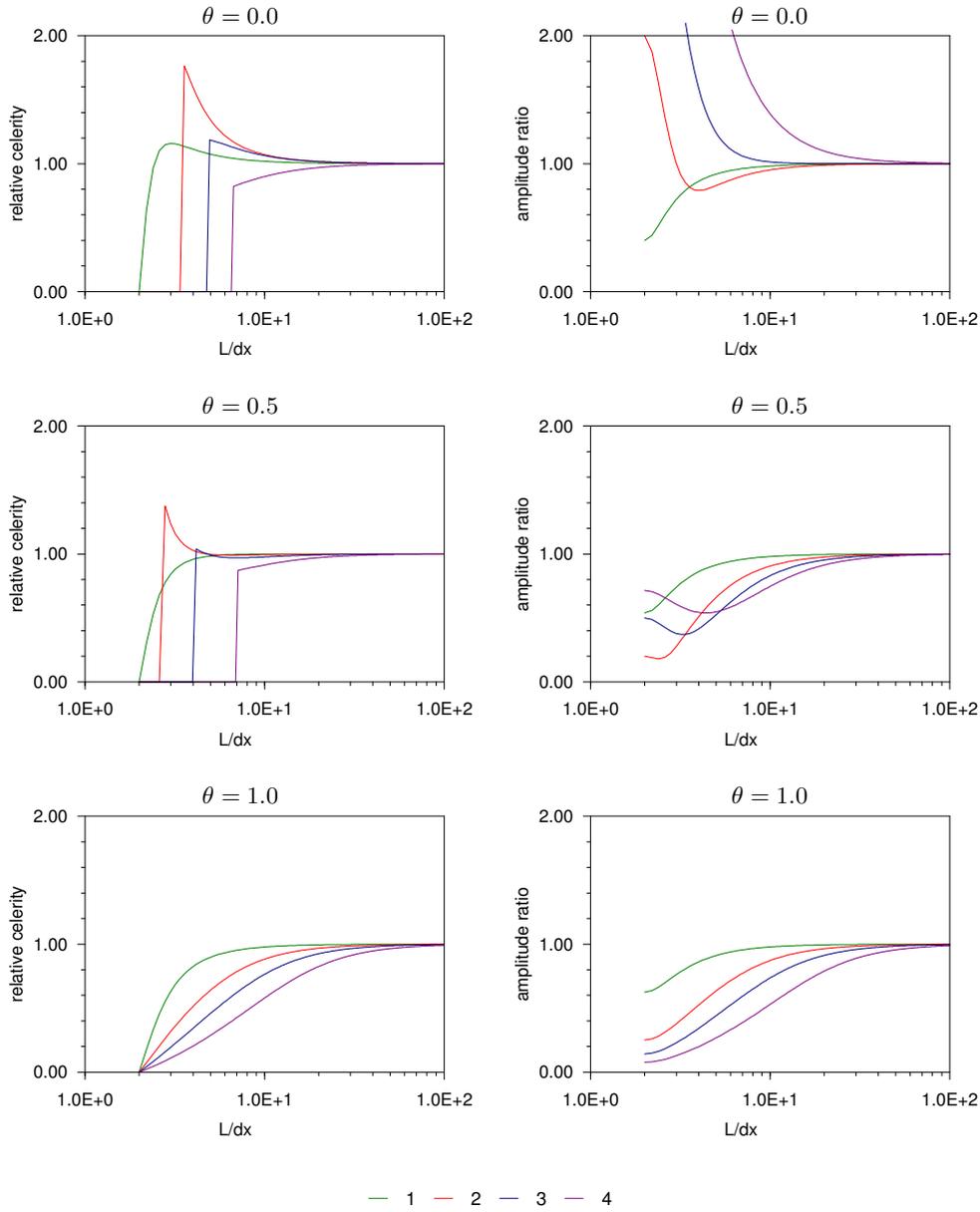


Figure 3.1: Amplification factors for consistent-mass formulation 1:  $\lambda = \sigma = 0.1$ , 2:  $\lambda = \sigma = 0.5$ , 3:  $\lambda = \sigma = 1.0$ , 4:  $\lambda = \sigma = 2.0$ .

This condition states two restrictions on the maximal time step. The first condition  $\lambda \leq 1$  states  $\Delta t \leq \Delta t_1 = C\Delta x^2/2D$ . The second condition  $\sigma^2 \leq \lambda$  poses  $\Delta t \leq \Delta t_2 = 2CD/U^2$ . This condition was given already by equation (3.97). If the condition of the

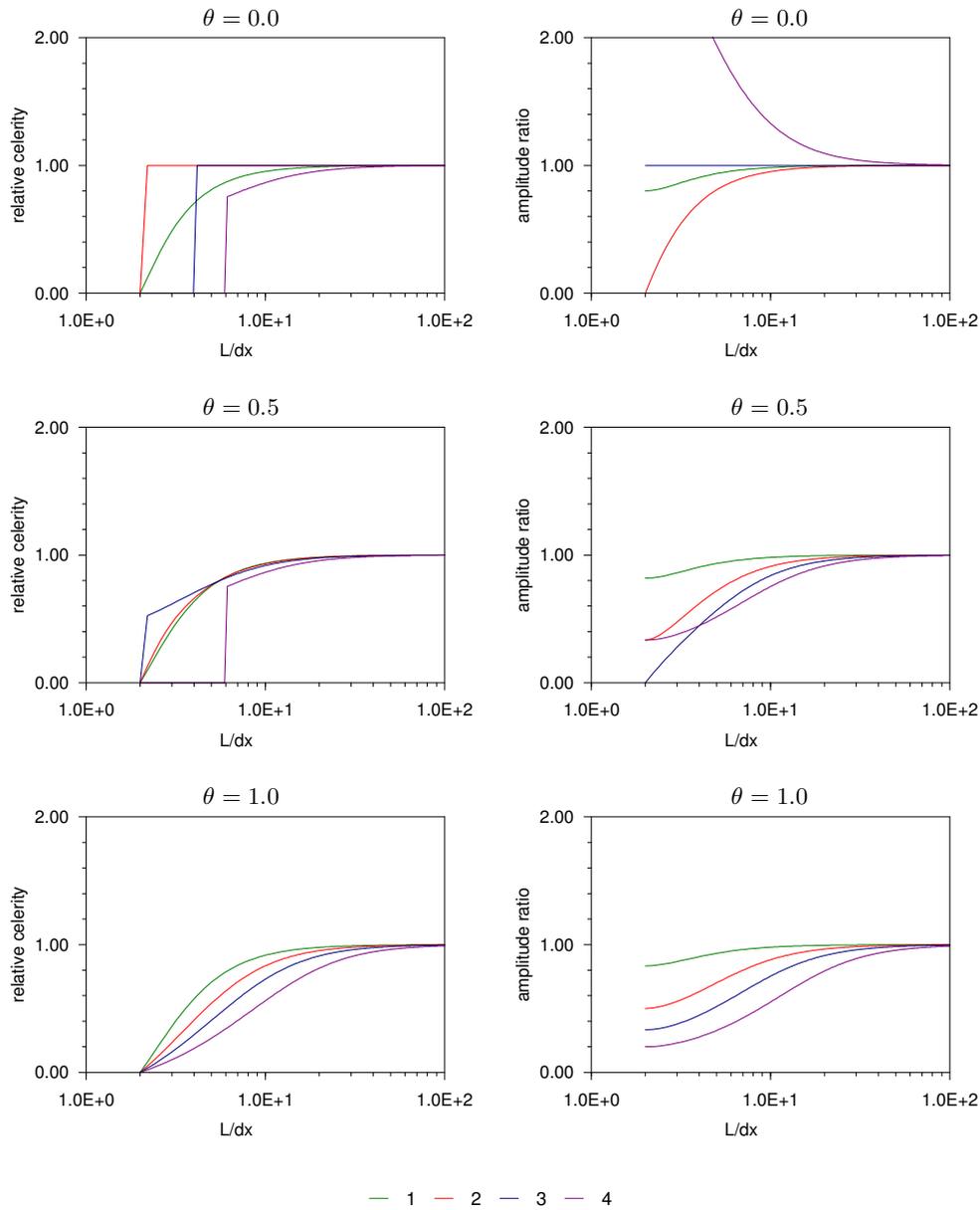


Figure 3.2: Amplification factors for lumped-mass formulation 1:  $\lambda = \sigma = 0.1$ , 2:  $\lambda = \sigma = 0.5$ , 3:  $\lambda = \sigma = 1.0$ , 4:  $\lambda = \sigma = 2.0$ .

cell Reynolds number or cell Peclet number given by  $Pe = |U\Delta x/2D| \leq 1$  holds, then the first condition is more restrictive as the Peclet number follows from  $\Delta t_1 \leq \Delta t_2$ . Equation (3.105) states that the Courant-Friedrichs-Lewy condition, expressed by the

Courant number  $\text{Cr} = |U\Delta t/C\Delta x|$  is necessary but not sufficient as  $\lambda \leq 1$  and  $\text{Cr} = \sigma$ . For pure convection the explicit scheme is unconditionally unstable. For pure diffusion only the first time step criterion holds. An alternative formulation of equation (3.105) reads

$$\text{Cr}^2 \leq \text{Cr}/\text{Pe} \leq 1. \quad (3.106)$$

If the spatial discretization follows from the cell Peclet number, then the Courant number should be smaller than the Peclet number if the Peclet number is less than unity. The Courant number should be less than the reciprocal value of the Peclet number if the Peclet number exceeds one. The time step size subsequently follows from the Courant number. In two-dimensional or three-dimensional problems the time step size depends on the cell with maximum velocity. In general, its position will vary in time.

For the flow problem the first stability criterion specifies  $\lambda$  by

$$\lambda^p = \frac{2k_r K_{max} \Delta t}{\mu \alpha_p S \Delta x_{min}^2}, \quad (3.107)$$

where  $K_{max}$  denotes the maximum eigenvalue of the intrinsic permeability tensor and  $\Delta x_{min}$  is approximated by the diameter of the incircle of the element. For an explicit formulation this number has to hold  $\lambda < 1$ . The cell Peclet number for the solute transport equation reads

$$\text{Pe}^\omega = \frac{\Delta x_{max}}{2\alpha_l}, \quad (3.108)$$

and the Courant number is given by

$$\text{Cr}^\omega = \frac{|q|\Delta t}{nS\Delta x_{min}}, \quad (3.109)$$

where the ratio of dispersivity tensor  $D_{ij}$  to Darcy's flux  $q_i$  equals the longitudinal dispersivity  $\alpha_l$ ,  $\Delta x_{max}$  denotes the diameter of the circumcircle. For the heat transport equation the Peclet number is given by

$$\text{Pe}^T = \frac{\rho c |q| \Delta x_{max}}{2(nS\lambda_l + (1-n)\lambda_l^s + c\rho\alpha_l q)}. \quad (3.110)$$

The Courant number for this equation reads

$$\text{Cr}^T = \frac{\rho c |q| \Delta t}{(nS\rho c + (1-n)\rho^s c^s) \Delta x_{min}}. \quad (3.111)$$

The stability limits given above have to be satisfied for the explicit formulation of the linearized basic equations. The conditions can also give some guidance for solving the non-linear flow and transport equation implicitly. Although the implicit time integration provided an unconditional stable method for solving the test problem, the time step is limited by the range of convergence of the nonlinear solver for more advanced problems.

The accuracy of the solution depends on the time step size and the size of the elements. On non-uniform grids the accuracy of the spatial discretization also depends

on the ratio of adjacent element sizes, as lower order Taylor terms do not cancel out. Comparing computational results with results obtained on a globally refined mesh and decreased time step size provides a more rigorous indication on the accuracy of the solution. Numerical diffusion produced by time integration, can effectively be reduced by decreasing the time step size. Wiggles also relate to the accuracy of the discretization method [119]. These short waves,  $L \approx 2\Delta x$ , may occur in both steady-state and transient convection-diffusion problems if the Peclet condition is not satisfied. The convective form of the transport equations sets the diffusive flux over the outflow boundary to zero. This prevents the boundary condition to trigger wiggles. Still permeability contrasts in the domain may generate regions with steep concentration gradients, which are able to trigger wiggles. Local refinement reduces the Peclet number and suppresses wiggles.

### 3.2.2 Initial conditions

Initial conditions are given by nodal values of the primary variables: pressure  $p_a$ , mass fraction  $\omega_a$  and temperature  $T_a$ . Basis functions interpolate the nodal values as  $p = p_a N_a$ ,  $\omega = \omega_a N_a$  and  $T = T_a N_a$  and span three continuous parameter fields. If these fields are specified separately, generally they do not comply with an equilibrium condition.

A better way to state the initial condition is to perform a steady state calculation for the coupled flow and transport problem, given a proper set of boundary conditions. The steady state solution can be obtained from the set of governing algebraic equations by setting the time derivatives to zero. Alternatively, a time step algorithm can be used as an iterative scheme to derive the steady state situation. The second approach gives best results in combination with adaptive mesh refinement. For both schemes the starting vectors for the primary variables need still to be specified. For this a constant mass fraction field specifies the mass fraction vector as

$$\omega = \omega^0 \quad \text{on } \Omega, \quad (3.112)$$

where  $\omega^0$  could be specified by the reference mass fraction  $\omega_0$ , which is part of the expression for the fluid density. The reference temperature then specifies the starting temperature field as

$$T = T^0 \quad \text{on } \Omega, \quad (3.113)$$

and a fully saturated hydrostatic pressure field specifies the initial pressure distribution. This field is given by

$$p = \rho g (z_{max} - z) \quad \text{on } \Omega, \quad (3.114)$$

where  $z$  is pointing in the opposite direction of the gravity vector, and  $z_{max}$  is larger than the length of the domain in  $z$ -direction to ensure a saturated situation in the entire flow domain. If the fluid density relates linearly with solute mass fraction and temperature, the density can be approximated by the reference density  $\rho_0$ , as it shows a weak relation to the pore pressure. In general, the density is calculated from the solute mass fraction field and the initial temperature field. Nonlinearities in the starting vector are disregarded. This situation suppresses nonlinearities associated with partly

unsaturated domains. Numerical tests show that the solution of a de-watering problem is easier obtained than the solution of a re-wetting problem.

### 3.2.3 Boundary conditions

No-flow boundary conditions of equation (2.92) pose natural conditions on the boundary segment  $\Gamma_1$  where they apply. The volumetric boundary flux reads  $\bar{q}_n = 0$  and the condition combines the previously defined condition as

$$\Gamma_1 = \Gamma_2^p \cap \Gamma_2^\omega \cap \Gamma_2^T. \quad (3.115)$$

Inflow boundary conditions may either be prescribed by a predefined volumetric boundary flux, given by equation (2.93), on the segment  $\Gamma_2$  or may result from a given pressure for equation (2.93) at the boundary  $\Gamma_3$ . The prescribed boundary flux  $\bar{q}_n > 0$  is imposed by a combination of

$$\Gamma_2 = \Gamma_2^p \cap \Gamma_2^\omega \cap \Gamma_2^T. \quad (3.116)$$

Inflow boundary conditions which result from a prescribed pressure  $\bar{p}$  combine

$$\Gamma_3 = \Gamma_1^p \cap \Gamma_2^\omega \cap \Gamma_2^T. \quad (3.117)$$

Outflow boundary conditions are also obtained by either a prescribed boundary flux, given by equation (2.95), or by a prescribed pressure as equation (2.96) on the boundary elements where they apply. Prescribed boundary fluxes on  $\Gamma_4$  for which ( $\bar{q}_n < 0$ ) combine the conditions on the basic equations as

$$\Gamma_4 = \Gamma_2^p \cap \Gamma_2^\omega \cap \Gamma_2^T. \quad (3.118)$$

Outflow boundary conditions on  $\Gamma_5$  generated by a predefined pore water pressure  $\bar{p}$ , gather the separately formulated conditions according to

$$\Gamma_5 = \Gamma_1^p \cap \Gamma_2^\omega \cap \Gamma_2^T. \quad (3.119)$$

Here the third-type boundary conditions for the transport equations do not contribute to the formulated set of equations as the boundary integrals only apply to second-type inflow boundary conditions.

Artificial conditions on boundaries that separate parts of the porous medium do not distinguish between inflow and outflow. Both inflow and outflow conditions are based on a predefined boundary flux  $\bar{q}_n$ , given by equation (2.97). The conditions are set as

$$\Gamma_6 = \Gamma_2^p \cap \Gamma_1^\omega \cap \Gamma_1^T. \quad (3.120)$$

Artificial inflow and outflow conditions which include a prescribed pressure condition  $\bar{p}$ , according to equation (2.98) are given by

$$\Gamma_7 = \Gamma_1^p \cap \Gamma_1^\omega \cap \Gamma_1^T. \quad (3.121)$$

The definition of geometrical nonlinear conditions depends on the imposed boundary flux. Seepage conditions of equation (2.99) specify  $\bar{q}_n = 0$ , precipitation conditions

given by equation (2.100) impose  $\bar{q}_n > 0$  and evaporation conditions of equation (2.101) propose  $\bar{q}_n < 0$ . Precipitation conditions and evaporation conditions both generalize the seepage condition. The boundary conditions impose either  $\Gamma_1^p$  or  $\Gamma_2^p$  conditions. Seepage conditions impose natural  $\Gamma_2^p$  conditions on the boundary segments if unsaturated pore pressures are calculated. If imposing these conditions results in saturated pore pressures then the boundary type changes to a  $\Gamma_1^p$ . These conditions should not generate an in-flux. The seepage boundary condition is written as

$$\begin{cases} p^{r+1} = 0 & \text{if } (\rho q_i)^r n_i \geq 0 \\ (\rho q_i)^{r+1} n_i = 0 & \text{for } p^r < 0 \end{cases}, \quad (3.122)$$

where  $r + 1$  denotes the next nonlinear iteration and  $r$  is the current iteration. First the condition of all seepage boundary nodes has to be checked and in the next iteration only the node with the largest error in the condition will be modified. Alternatively the node with the smallest elevation could be updated. In general, bad convergence behavior results from updating all nodes at the same iteration step.

Precipitation boundaries replace the atmospheric zero pressure condition by a ponding depth pressure  $p = \bar{p}_p$  and replace the condition on the flux by  $q_n = \bar{q}_n$  corresponding to the precipitation flux. The predefined flux has a positive value and the precipitation condition reads

$$\begin{cases} p^{r+1} = \bar{p}_p & \text{if } -(\rho q_i)^r n_i \leq \rho \bar{q}_n \\ -(\rho q_i)^{r+1} n_i = \rho \bar{q}_n & \text{for } p^r < \bar{p}_p \end{cases}. \quad (3.123)$$

Evaporation conditions also generalize seepage conditions. The evaporation condition replaces the atmospheric pressure limit by a pressure based on the humidity of air  $p = \bar{p}_e$  and replace the condition on the flux by a potential evaporation flux  $q_n = \bar{q}_n$ .

$$\begin{cases} p^{r+1} = \bar{p}_e & \text{if } (\rho q_i)^r n_i \leq -\rho \bar{q}_n \\ (\rho q_i)^{r+1} n_i = -\rho \bar{q}_n & \text{for } p^r > \bar{p}_e \end{cases}. \quad (3.124)$$

The relative permeability of a drying top soil will decrease and after a while the soil will not be able to produce the potential outflux. The actual flux then follows from the pressure condition.

A rising surface water level overrules the geometrical nonlinear conditions if the pressure is higher than the ponding depth pressure. A functional relation of the water level in time is imposed by

$$p^n = \bar{p}(t). \quad (3.125)$$

If this condition applies, inflow or outflow could occur.

# Chapter 4

## Adaptive multiscale method

Adaptive grid refinement aims to reach fine-scale accuracy on partially refined grids. It reduces the size of the system of algebraic equations, and cuts down the computational costs for solving the set of equations. Compared to a fine-scale discretization, the number of integration points is less, which saves the computational work in setting up the system of equations. Coarsening the mesh on a heterogeneous field requires the computation of equivalent field behavior. Multiscale finite element methods bring fine scale information to the coarse scale by calculating multiscale basis functions [24, 2, 63, 68, 69, 73, 80]. The finite element procedure then constructs the algebraic coarse scale equations by using these functions. Subsequently the global flow problem is solved at the coarse mesh. The multiscale basis functions extrapolate the coarse scale solution to the fine scale. In order to obtain the multiscale basis functions the domain is decomposed in coarse scale elements, and flow problems are solved locally. The local problems can be closed in their subdomains by the solution of dimensionally reduced flow problems. This closure introduces errors in the global solution if flow on the coarse scale is not tangential to the domain boundary. Unlike other domain decomposition techniques, this error is not corrected iteratively [98]. The proposed adaptive multiscale finite element method corrects these errors by refining the mesh locally in zones of interest. The method is attractive when applying parallel computing. The formulation facilitates a multigrid solver. Intergrid transfer operators interpolate fine-scale information to the coarse scale and extrapolate coarse scale information back to the fine scale. Coarse grid operators apply to the coarsened meshes.

Section 4.1 presents the adaptive formulation. Discrete adaptivity refines the finite element mesh locally on homogeneous and piecewise homogeneous fields. Functional adaptivity supports coarsening of the mesh on homogeneous fields. Section 4.2 presents the multiscale formulation by introducing multiscale basis functions and constructing the coarse mesh discretization. Section 4.3 outlines the multigrid formulation. The multiscale formulations will be illustrated on a number of periodic fields, further elaborated in appendix B. Appendix C introduces the multigrid method that will be used to solve the algebraic equations, both globally and locally.

## 4.1 Adaptive formulation

Enrichment of finite element meshes may either be predefined, or may depend on calculated results. Static refinement increases the accuracy for instance around wells and near tips of sheet pile walls or resolves the resolution of a highly permeable streak. For static refinement the refined parts of the mesh do not change in time. Dynamic refinement is able to capture for instance concentration fronts. For dynamic mesh refinement the refined areas do change in time. Adaptive refinement requires the postulation of a set of refinement criteria. Figure 4.1 illustrates an adaptive mesh refinement procedure on four finite element levels.

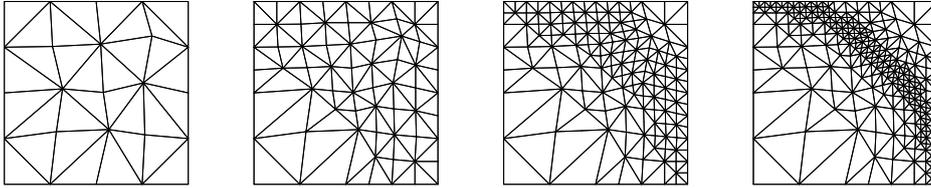


Figure 4.1: Mesh refinement.

Section 4.1.1 focuses on the adaptive formulation, and presents a nested hierarchy of finite element meshes. It introduces variational coarsening, which merges fine scale elements into a single coarse scale element, and regularization, which obtains a gradual refinement. At the end of the section a refinement criterion is postulated. Section 4.1.2 focuses on functional adaptivity. It introduces a numerical homogenization algorithm called pressure-dissipation averaging for which local boundary conditions have to be proposed. Section 4.3 outlines the multigrid solver, which applies the multiscale formulation to speed up its convergence behavior.

### 4.1.1 Discrete adaptivity

In order to facilitate the multiscale algorithm, the adapted mesh is composed of parts of a nested hierarchy of meshes. Figure 4.2 shows a mesh hierarchy with uniform mesh refinement. The proposed adaptive formulation use simplex elements, which are more flexible in dealing with complex geometries. Simplex elements, like triangular and tetrahedral elements construct conforming meshes without hanging grid points more generally, which is beneficial for the accuracy of an adaptive method [112, 113].

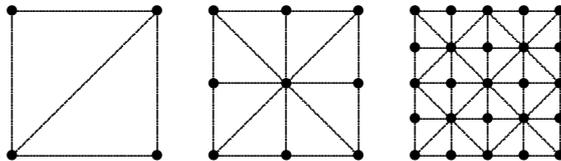


Figure 4.2: Constant mesh hierarchy.

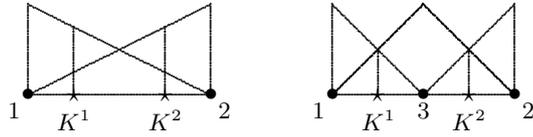


Figure 4.3: Standard basis functions for line elements.

### Element hierarchies

Figure 4.3 illustrates the hierarchical formulation of finite element basis functions. The left picture shows the layout of a test problem on the domain  $x \in [0, 1]$  and standard basis functions for a line element, the right picture shows the standard basis functions for its child elements.

The standard basis functions for the test problem, on the coarse grid, in global coordinates, read

$$N_1^{k-1} = 1 - x, \quad N_2^{k-1} = x. \quad (4.1)$$

Standard basis functions on the fine-grid for the first element are given by

$$N_1^k = 1 - 2x, \quad N_3^k = 2x, \quad (4.2)$$

and for the second element they read

$$N_2^k = 2x - 1, \quad N_3^k = 2 - 2x. \quad (4.3)$$

Standard coarse-scale basis functions  $N_a^{k-1}$  for a line element can be expressed as the sum of its child elements basis functions  $N_a^k$  by

$$N_1^{k-1} = N_1^k + \frac{1}{2}N_3^k, \quad N_2^{k-1} = N_2^k + \frac{1}{2}N_3^k. \quad (4.4)$$

A generalized formulation in index notation reads

$$N_a^{k-1} = P_{ba}^k N_b^k. \quad (4.5)$$

The standard interpolation matrix for line elements  $[P^k]$  then reads

$$[P^k] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1/2 & 1/2 \end{bmatrix}. \quad (4.6)$$

The coarsening procedure will be given for the one-dimensional test problem illustrated by figure 4.3. This problem considers steady state saturated groundwater flow with constant fluid properties on the flow domain  $x \in [0, 1]$ . The gravity vector points in opposite x-direction. The permeability shows a discontinuity for  $x = 0.5$  and has a value  $K_1$  for  $x \in [0, 0.5]$  and a value  $K_2$  for  $x \in [0.5, 1]$ . The problem will be solved by two line elements on the fine scale, where the fine-scale node is positioned in the middle of both coarse-scale nodes.

Matrix equations representing steady state groundwater flow at a coarse scale  $k-1$  written as  $L_{ab}^{k-1}u_b^{k-1} = f_a^{k-1}$ , can be derived by using basis functions or their derivatives at the coarse level. The consistent (coarse-grid) approach applies a weighted summation over fine-scale integration points, and computes matrix components as

$$L_{ab}^{k-1} = \sum_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho k_r K}{\mu} \frac{\partial N_a^{k-1}}{\partial x} \frac{\partial N_b^{k-1}}{\partial x} d\Omega, \quad (4.7)$$

approximated by

$$L_{ab}^{k-1} \approx \sum_{e=1}^{n_e^{k-1}} \tilde{w}_m^e \left( \frac{\rho k_r K}{\mu} \frac{\partial N_a^{k-1}}{\partial \xi} \frac{\partial \xi}{\partial x} \frac{\partial N_b^{k-1}}{\partial \xi} \frac{\partial \xi}{\partial x} |J| \right)_m^e, \quad (4.8)$$

where  $\rho$  denotes the mass density of the fluid [ $\text{ml}^{-3}$ ],  $k_r$  represents the relative permeability of the porous medium  $[-]$ ,  $K$  is the intrinsic permeability of the solid matrix [ $\text{l}^2$ ] and  $\mu$  expresses the dynamic viscosity of the fluid [ $\text{ml}^{-1}\text{t}^{-1}$ ]. The weights at fine-scale integration points  $m$  are given by  $\tilde{w}_m^e$ . The points are located at fine grid positions. The superscript  $e$  denotes the element number and  $n_e^{k-1}$  indicates the number of elements that compose the coarse-scale mesh. Inserting the basis functions given by equation (4.1) gives a coarse-grid matrix represented by

$$[L^{k-1}] = \frac{\rho k_r}{2\mu} \begin{bmatrix} K^1 + K^2 & -K^1 - K^2 \\ -K^1 - K^2 & K^1 + K^2 \end{bmatrix}, \quad (4.9)$$

where  $K^1$  is the intrinsic permeability measured in the first integration points, which is located in the first fine-scale element and  $K^2$  is attached to the second fine-scale element. It's value is obtained in the second integration point.

The right hand side of the discrete flow equation is written as

$$f_a^{k-1} = \sum_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho k_r K}{\mu} \frac{\partial N_a^{k-1}}{\partial x} \rho g d\Omega \quad (4.10)$$

or

$$f_a^{k-1} \approx \sum_{e=1}^{n_e^{k-1}} w_m^e \left( \frac{\rho k_r K}{\mu} \frac{\partial N_a^{k-1}}{\partial \xi} \frac{\partial \xi}{\partial x} \rho g |J| \right)_m^e, \quad (4.11)$$

where  $g$  denotes the gravitational force [ $\text{lt}^{-2}$ ],  $g$  has a negative value as the gravity vector points in the opposite direction to the local coordinate vector. Coarse-scale basis functions given by equation (4.1) yield

$$[f^{k-1}] = \rho g \frac{\rho k_r}{2\mu} \begin{bmatrix} -K^1 - K^2 \\ K^1 + K^2 \end{bmatrix}. \quad (4.12)$$

The coarse-grid matrix and vector could alternatively be constructed by assembling the fine-scale matrix over the fine-grid element domains that contribute to the coarse parent element using the standard interpolation matrix given by equation (4.6) in a

variational (fine-grid) approach. This formulation expresses the coarse scale matrix components as

$$L_{ab}^{k-1} = \sum_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho^{k_r} K}{\mu} P_{ca}^k \frac{\partial N_c^k}{\partial x} P_{db}^k \frac{\partial N_d^k}{\partial x} d\Omega, \quad (4.13)$$

approximated by

$$L_{ab}^{k-1} \approx \sum_{e=1}^{n_e^k} w_m^e \left( \frac{\rho^{k_r} K}{\mu} P_{ca}^k \frac{\partial N_c^k}{\partial \xi} \frac{\partial \xi}{\partial x} P_{db}^k \frac{\partial N_d^k}{\partial \xi} \frac{\partial \xi}{\partial x} |J| \right)_m^e, \quad (4.14)$$

and the right hand side vector components as

$$f_a^{k-1} = \sum_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho^{k_r} K}{\mu} P_{ca}^k \frac{\partial N_c^k}{\partial x} \rho g d\Omega, \quad (4.15)$$

or

$$f_a^{k-1} \approx \sum_{e=1}^{n_e^k} w_m^e \left( \frac{\rho^{k_r} K}{\mu} P_{ca}^k \frac{\partial N_c^k}{\partial \xi} \frac{\partial \xi}{\partial x} \rho g |J| \right)_m^e. \quad (4.16)$$

The components of the interpolation matrix  $P_{ab}^k$  construct the matrix  $[P^k]$ . The weights of the fine-scale integration points are given by  $w_m^k = 1$ , as summation over a single integration point ( $n_m = 1$ ) provides the exact value of the integral for constant porous media properties. Fine-scale basis functions given by equations (4.2) and (4.3) and standard interpolation given by equation (4.6) also produce the coarse-grid matrix of equation (4.9) and coarse-grid vector of equation (4.12). Inconsistent results will generally be obtained if the integration points in the coarse-grid elements do not coincide with the integration points of a single fine-grid element for the case of a non-homogeneous field function  $K$ , which varies over the fine scale elements.

In this thesis, the coarsening process of a multi-dimensional problem is carried out sequentially. The procedure facilitates two fine scale elements to merge into a single coarse element. For the triangular child elements of figure 4.4 this procedure is carried

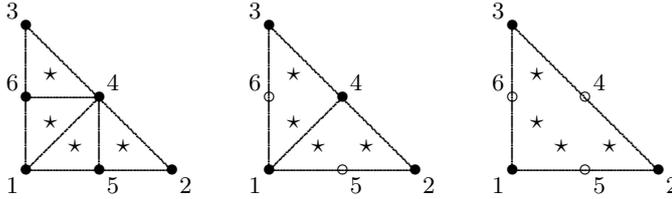


Figure 4.4: Multistep coarsening of triangular elements.

out in two steps. In the first step, element (6, 4, 3) and element (1, 4, 6) are mapped and generate the new element (1, 4, 3). Element (5, 2, 4) and element (1, 5, 4) produce the new element (1, 2, 4). In the second step element (1, 4, 3) and element (1, 2, 4) merge into the coarse-grid element (1, 2, 3). Figure 4.5 depicts the hierarchy of parent and child elements. Variational coarsening merges basis functions of fine scale elements into coarse scale basis functions of their parent element, and removes hanging grid nodes.

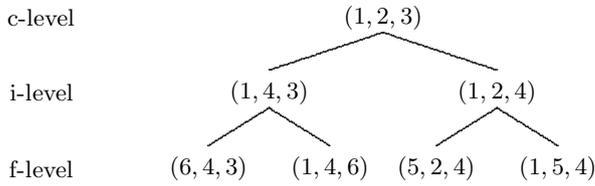


Figure 4.5: Triangular element hierarchy.

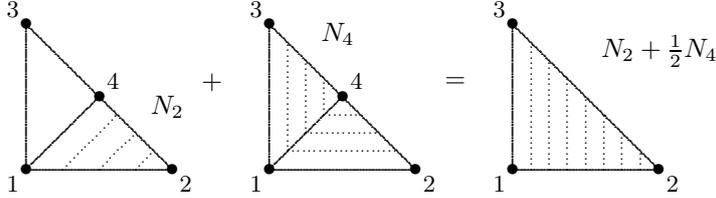


Figure 4.6: Merged fine element basis functions.

Figure 4.6 illustrates the merging procedure over two triangular elements. The coarse scale triangular basis functions in global coordinates read

$$N_1^{k-1} = 1 - x - y, \quad N_2^{k-1} = x, \quad N_3^{k-1} = y. \quad (4.17)$$

Fine scale basis functions over element (1,2,4) are written as

$$N_1^k = 1 - x - y, \quad N_2^k = x, \quad N_3^k = 0, \quad N_4^k = y, \quad (4.18)$$

and fine scale basis functions over element (1,3,4) read

$$N_1^k = 0, \quad N_2^k = 1 - y, \quad N_3^k = -1 + x + y, \quad N_4^k = 1 - x. \quad (4.19)$$

Linear interpolation reproduces standard basis functions on the coarse scale, which can be written as

$$N_1^{k-1} = N_1^k, \quad N_2^{k-1} = N_2^k + \frac{1}{2}N_4^k, \quad N_3^{k-1} = N_3^k + \frac{1}{2}N_4^k. \quad (4.20)$$

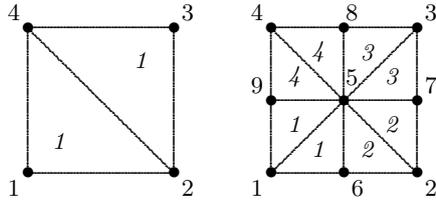


Figure 4.7: Triangular parent and child element patches.

In order to obtain equivalent material behavior via the computation of multiscale basis functions, as will be explained in the next section, the standard basis functions

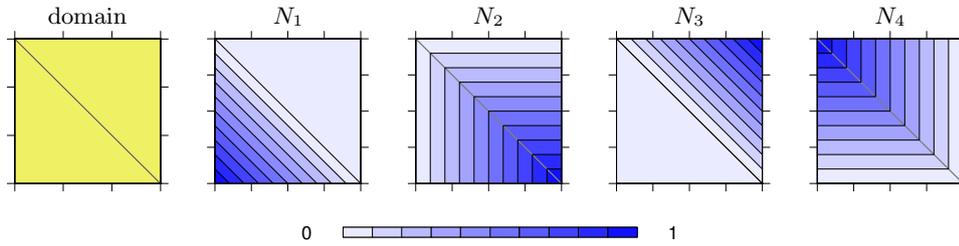


Figure 4.8: Linear basis functions over a span of two triangular elements.

over a patch of elements will be formulated. Figure 4.7 shows a patch of triangular elements, which forms a quadrilateral domain. Coarse scale basis functions over the patch of triangular elements follow from the fine scale basis functions as

$$\begin{aligned}
 N_1^{k-1} &= N_1^k + \frac{1}{2}N_6^k + \frac{1}{2}N_9^k, \\
 N_2^{k-1} &= N_2^k + \frac{1}{2}N_6^k + \frac{1}{2}N_7^k + \frac{1}{2}N_5^k, \\
 N_3^{k-1} &= N_3^k + \frac{1}{2}N_7^k + \frac{1}{2}N_8^k + \frac{1}{2}N_5^k, \\
 N_4^{k-1} &= N_4^k + \frac{1}{2}N_8^k + \frac{1}{2}N_9^k.
 \end{aligned} \tag{4.21}$$

Figure 4.8 presents linear basis functions on two triangular elements, which constitute to the patch, graphically.

### Refinement procedure

Figure 4.9 depicts a mesh hierarchy with local mesh refinement. The finest grid in this figure shows non-conforming elements and hanging nodes indicated by the symbol  $\diamond$ . The shape functions across the non-conforming element boundaries are in general not continuous. For this case continuous shape functions are obtained by merging of elements. For more general situations, adaptive refinement may still result in non-

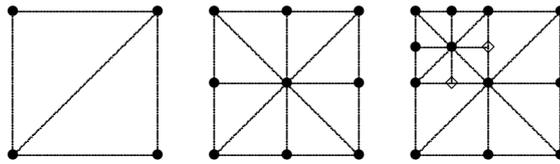


Figure 4.9: Adaptive mesh hierarchy.

conforming meshes if hanging grid points occur on multiple grid levels. To circumvent this a regularization procedure is proposed here. Figure 4.10 illustrates this procedure. The upper four pictures of figure 4.10 show the process of adaptive refinement for a triangular element mesh without regularization. The first refinement step generates four child elements on level 2 for the lower left element. The next step refines one level 2 element and the last step refines two level 3 elements. The mesh on the finest level

consists of 14 elements. In the final configuration, one of the level 2 elements is attached to two level 4 elements and one level 3 element. On the internal boundary hanging nodes are present. The four pictures at the bottom of figure 4.10 show the hierarchy of meshes when adopting mesh regularization. The first refinement step already produces a hanging grid node that can not be removed by merging level 2 elements. Here, the adjacent coarse element needs to be refined as well. However, the four child elements of this element can merge into two transition elements. The second refinement step generates four level 3 elements, two of them could form a transition element making renormalization unnecessary. The last refinement step introduces 16 level 4 elements and eight of them merge into 4 transition elements. The most right transition elements are connected to the hanging grid node, which was generated in the second refinement step, and now the adjacent level 2 elements have to be refined. This requirement also introduces four extra level 3 elements and two merged level 3 elements.

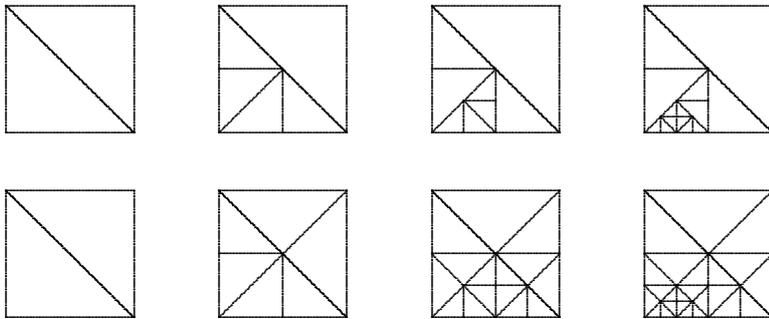


Figure 4.10: Adaptive mesh refinement without and with regularization.

The regularization algorithm obtains its result by setting a counter for each node equal to the finest level of the elements that share this node. Comparing the element levels with their nodal counters produces flags for elements that need to be refined. If the difference between levels and counters is larger than one, then the element needs to be refined and the level for these elements has to be increased by one. This procedure is repeated until no element level has to be increased. After identifying all elements that need to be refined, the final mesh can be constructed, without the need to repeat the regularization.

### Refinement criteria

Conformal finite elements obtain nodal based mass continuity. The pressure gradient tangential to the element boundary is continuous, whereas the gradient perpendicular to the boundary is discontinuous and continuity of flux over the element boundary is not preserved. The refinement criterion proposed here, detects this discontinuity and mesh enrichment corrects the accuracy of the flow field locally. This procedure improves the accuracy of the transport simulation, which is based on the velocity field. Figure 4.11 shows the Darcy flux vector in a single integration point and its components

perpendicular and tangential to the element boundaries in the flow points.

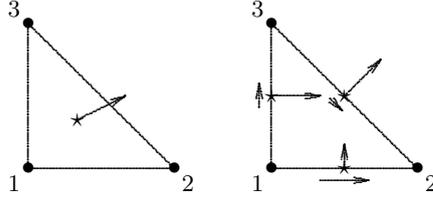


Figure 4.11: Darcy flux in integration point and flow points.

The expression for the Darcy mass flux components, postulated in global coordinates, and obtained in the flow points reads

$$q_i = \frac{\rho k_r K_{ij}}{\mu} \left( \frac{\partial N_c}{\partial \xi} \frac{\partial \xi}{\partial x_i} p_c - \rho g_i \right). \quad (4.22)$$

The flow points are located on the edges of two-dimensional elements and on the faces of three-dimensional elements. The magnitude of the volumetric flux over the element boundaries reads

$$q_n = q_i n_i, \quad (4.23)$$

where  $n_i$  is the outward pointing normal vector on the element edge. The difference in magnitude of the flux perpendicular to the element boundary, denoted by  $\Delta q_n$ , computes the refinement criterion. This criterion reads

$$\Delta q_n > c_1 \langle q^l \rangle, \quad \Delta q_n = |q_n^1 - q_n^2|, \quad (4.24)$$

where  $\langle q^l \rangle$  expresses the mean perpendicular flux over the global flow field.

Figure 4.12 illustrates the refinement procedure on a checkerboard cell. The cell is closed for flow on the horizontal boundary parts of the cell and prescribed constant pressure boundary conditions apply on the vertical boundary parts. The mesh refinement concentrates on the singular point in the center of the cell. The upper plots of figure 4.12 show three levels of full mesh refinement and one adaptively refined mesh. The lower plots relate to the upper pictures and show the pressure fields, which were obtained by three levels of fixed mesh refinement and one adaptive mesh refinement. The refinement procedure reaches fine scale accuracy on an adaptively refined mesh.

The accuracy of a numerical computation on a piecewise homogeneous field also depends on the resolution of the structure. Figure 4.13 poses the structural multiscale problem for diagonally layered cells. Adaptivity captures the geometry of the structures more accurately [115]. The upper plots of figure 4.13 illustrate the capability of structural adaptivity. The first three pictures resolve the structure by full mesh refinement, the last picture applies adaptive refinement and obtains the finest scale resolution on a partly refined grid. Adaptive refinement improves the accuracy of the pressure field as can be seen in the related pictures at the bottom in the figure. The boundary conditions prescribe a constant pressure at the vertical boundaries and a no-flow condition at the horizontal boundaries of the cell.

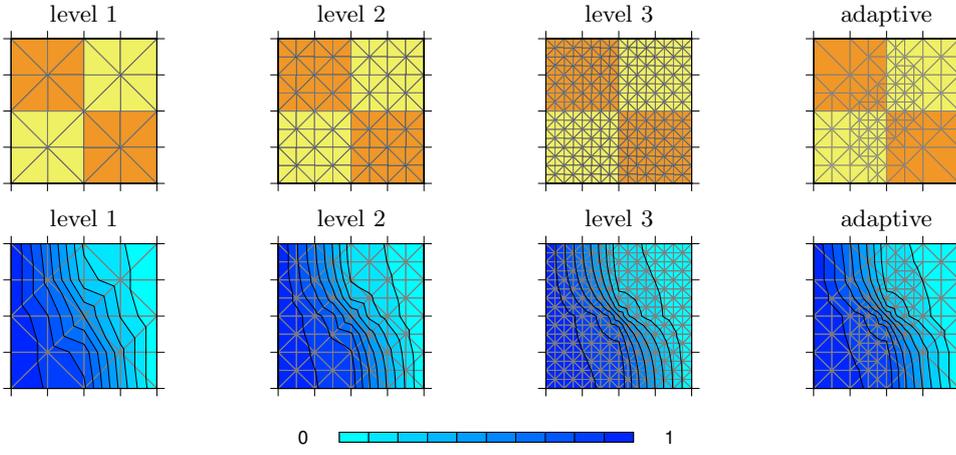


Figure 4.12: Discrete adaptivity on a checkerboard cell.

### 4.1.2 Functional adaptivity

Homogenization techniques are used to reduce the fine scale information by calculating equivalent material behavior on a coarse scale. Only if the scale of variation is much smaller than the coarse observation scale then the equivalent parameter is a constant variable on the sampling domain. In general, the equivalent parameter depends on the scale of the averaging domain and in the case of permeability, for instance, on closure of the domain. If the equivalent parameters are computed numerically, then their values also depend on the discretization of the flow problem. Refining the mesh increases the accuracy of the solution, and reduces the effect of local boundary conditions. Homogenization techniques obtain equivalent permeabilities either by solving local flow problems or by solving the global flow problem a limited number of times. Alternatively, the local prediction could be updated during the actual global flow simulation.

#### Variational averaging

For the one-dimensional test case given in section 4.1.1 the effective permeability  $\langle K \rangle$  could be sampled in the coarse-scale integration point. Elaboration gives

$$L_{ab}^{k-1} = \mathop{\text{A}}_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho k_r \langle K \rangle}{\mu} \frac{\partial N_a^{k-1}}{\partial x} \frac{\partial N_b^{k-1}}{\partial x} d\Omega, \quad (4.25)$$

approximated by

$$L_{ab}^{k-1} \approx \mathop{\text{A}}_{e=1}^{n_e^{k-1}} w_m^e \left( \frac{\rho k_r \langle K \rangle}{\mu} \frac{\partial N_a^{k-1}}{\partial \xi} \frac{\partial \xi}{\partial x} \frac{\partial N_b^{k-1}}{\partial \xi} \frac{\partial \xi}{\partial x} |J| \right)_m^e, \quad (4.26)$$

where  $w_m^{k-1} = 1$  for a single integration point. The coarse-scale matrix then reads

$$[L^{k-1}] = \frac{\rho k_r}{\mu} \begin{bmatrix} \langle K \rangle & -\langle K \rangle \\ -\langle K \rangle & \langle K \rangle \end{bmatrix}. \quad (4.27)$$

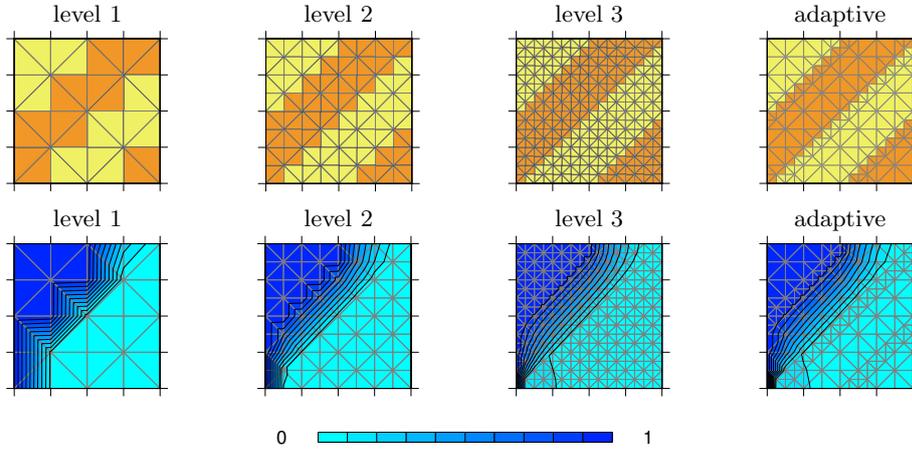


Figure 4.13: Functional adaptivity for diagonal layered cells.

The right hand side of the discrete flow equation is written as

$$f_a^{k-1} = \mathbf{A}_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho k_r \langle K \rangle}{\mu} \frac{\partial N_a^{k-1}}{\partial x} \rho g d\Omega, \quad (4.28)$$

or

$$f_a^{k-1} \approx \mathbf{A}_{e=1}^{n_e^{k-1}} w_m^e \left( \frac{\rho k_r \langle K \rangle}{\mu} \frac{\partial N_a^{k-1}}{\partial \xi} \frac{\partial \xi}{\partial x} \rho g |J| \right)_m^e. \quad (4.29)$$

The coarse scale vector reads

$$[f^{k-1}] = \rho g \frac{\rho k_r}{2\mu} \begin{bmatrix} -\langle K \rangle \\ \langle K \rangle \end{bmatrix}. \quad (4.30)$$

Equations (4.9) and (4.12) are retrieved, if the arithmetic averaged (equivalent) permeability  $\langle K \rangle = (K_1^2 + K_2^2)/2$  is applied. This equivalent permeability does not match the effective permeability given by the harmonic mean  $\langle K \rangle = (2K_1^2 K_2^2)/(K_1^2 + K_2^2)$ . For an arithmetic average, high conductivities dominate the coarse-grid behavior. The proposed adaptive formulation applies a more advanced numerical homogenization technique called pressure-dissipation averaging [129].

### Pressure-dissipation averaging

The pressure-dissipation averaging approach presented in appendix A combines conservation of driving force and conservation of energy dissipation. The set of algebraic equations for this approach, in Cartesian coordinates, reads

$$\left\langle \frac{-\partial p^{(a)k}}{\partial x_i} \right\rangle \left\langle \frac{-\partial p^{(b)k}}{\partial x_j} \right\rangle \tilde{K}_{ij} = \left\langle -\frac{\partial p^{(a)k}}{\partial x_i} q_i^{(b)k} \right\rangle, \quad (4.31)$$

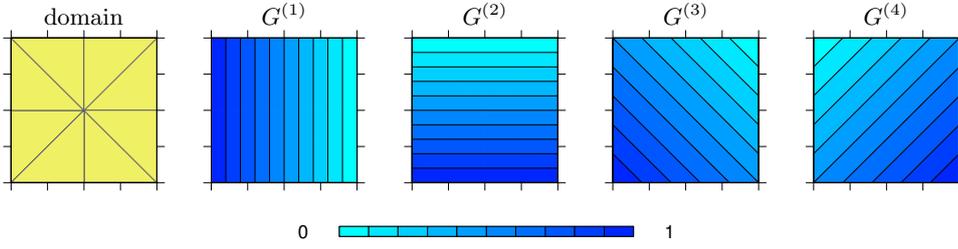


Figure 4.14: Linear boundary conditions on homogeneous (an)isotropic field.

where both (a) and (b) indicate loading cases. The flow problem on the patch could be closed by linear boundary conditions.

The linear loading cases read

$$G^{(1)} = 1 - x, \quad G^{(2)} = 1 - y, \quad (4.32)$$

or alternatively

$$G^{(3)} = 1 - \frac{1}{2}(x + y), \quad G^{(4)} = \frac{1}{2}(1 + x - y). \quad (4.33)$$

Figure 4.14 shows the pressure fields, which result from loading by linear boundary conditions on a homogeneous (an)isotropic cell. Two orthogonal loading cases give a set of four algebraic equations and solve a two-dimensional problem. The number of unknowns equals the square of the spatial dimension. For symmetric permeability tensors the number of unknowns reduces to three in two dimensions and six in three spatial dimensions. If off-diagonal components are disregarded, the number of unknowns equals the number of spatial dimensions  $n_d$ . Better known algorithms like pressure-flux averaging (appendix A) operate on a single loading case. This formulation produces  $n_d$  equations, which equals the number of diagonal tensor components. However, loading in x-direction produces flow in x-direction only, so  $K_{xx}$  can be calculated, and  $K_{yy}$  remains unknown. For both procedures  $n_d$  (preferably orthogonal) loading cases are required, to produce a non-singular matrix-vector equation. The set of algebraic equations, which follows from discretizing the mathematical pressure-dissipation problem is written as

$$\Phi_{ij}^{(ab)} \tilde{K}_{ij} = \phi^{(ab)}. \quad (4.34)$$

The fine-grid vector components are written as

$$\phi^{(ab)} = \mathbf{A}_{e=1}^{n_e^k} w_m^e \left( p_c^{(a)k} K_{ij} \frac{\rho k_r}{\mu} \frac{\partial N_c^k}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_d^k}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} p_d^{(b)k} |J| \right)_m^e. \quad (4.35)$$

The coarse-scale matrix components are given by

$$\Phi_{ij}^{(ab)} = \mathbf{A}_{e=1}^{n_e^{k-1}} w_m^e \left( p_c^{(a)k-1} \frac{\rho k_r}{\mu} \frac{\partial N_c^{k-1}}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_d^{k-1}}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} p_d^{(b)k-1} |J| \right)_m^e. \quad (4.36)$$

Load vector components  $p^{(a)k-1}$  and  $p^{(b)k-1}$  are given for the coarse scale and will be prolonged to the fine scale in the next section. The full matrix-vector form of the set

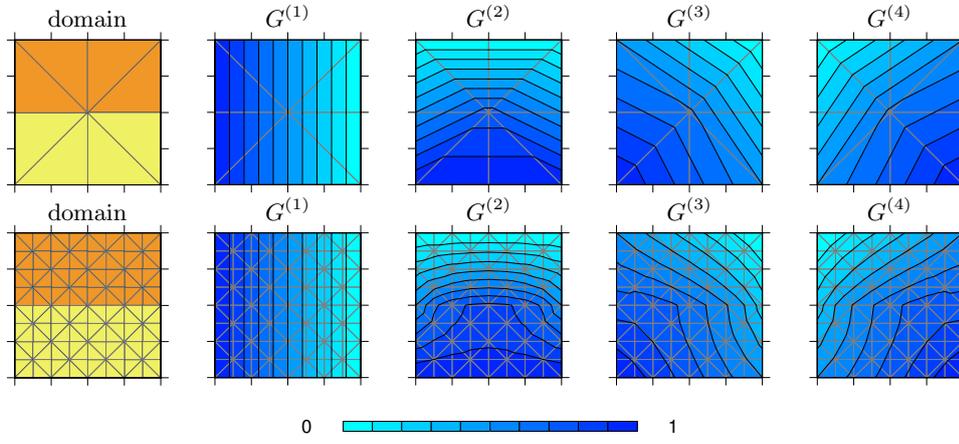


Figure 4.15: Linear boundary conditions on a horizontally layered cell.

of equations in two dimensions reads

$$\begin{bmatrix} \Phi_{xx}^{(11)} & \Phi_{xy}^{(11)} & \Phi_{yx}^{(11)} & \Phi_{yy}^{(11)} \\ \Phi_{xx}^{(12)} & \Phi_{xy}^{(12)} & \Phi_{yx}^{(12)} & \Phi_{yy}^{(12)} \\ \Phi_{xx}^{(21)} & \Phi_{xy}^{(21)} & \Phi_{yx}^{(21)} & \Phi_{yy}^{(21)} \\ \Phi_{xx}^{(22)} & \Phi_{xy}^{(22)} & \Phi_{yx}^{(22)} & \Phi_{yy}^{(22)} \end{bmatrix} \begin{bmatrix} \tilde{K}_{xx} \\ \tilde{K}_{xy} \\ \tilde{K}_{yx} \\ \tilde{K}_{yy} \end{bmatrix} = \begin{bmatrix} \phi^{(11)} \\ \phi^{(12)} \\ \phi^{(21)} \\ \phi^{(22)} \end{bmatrix}. \quad (4.37)$$

This expression produces a symmetric permeability tensor. An explicit expression follows from loading in x-direction and y-direction.

The set of equations will be solved for local element patches. Figure 4.7 presented the two-dimensional scaling patch. Fine-grid permeability tensors apply to the fine-grid patches (1, 6, 5, 9), (2, 7, 5, 6), (3, 8, 5, 7) and (4, 9, 5, 8), and equivalent anisotropic permeability coefficients are calculated for the coarse-grid patch (1, 2, 3, 4).

The pressure fields that result from loading by linear boundary conditions on a horizontally layered cell are presented by figure 4.15. Fine scale permeability tensors are considered to be isotropic. The intrinsic permeability of the low permeability material equals  $k = 10^{-2}$  and its value for the high permeability material is  $k = 1$ . The effective permeability components for this periodic cell read  $k_{xx} = 5.05 \cdot 10^{-1}$ ,  $k_{xy} = 0$ ,  $k_{yx} = 0$ , and  $k_{yy} = 1.98 \cdot 10^{-2}$ . The upper plots of the figure show fully-resolved finite element mesh solutions, and the lower plots give over-resolved solutions. The pictures show that mesh refinement increases the accuracy of the solution. Pressure-dissipation averaging extracts inaccurate permeability components from the fully-resolved pressure fields ( $k_{xx} = 5.05 \cdot 10^{-1}$ ,  $k_{xy} = -2.62 \cdot 10^{-17}$ ,  $k_{yx} = 2.78 \cdot 10^{-17}$ , and  $k_{yy} = 3.84 \cdot 10^{-1}$ ) and a bit more accurate results from the over-resolved pressure fields ( $k_{xx} = 5.05 \cdot 10^{-1}$ ,  $k_{xy} = -8.06 \cdot 10^{-17}$ ,  $k_{yx} = 6.93 \cdot 10^{-16}$ , and  $k_{yy} = 2.77 \cdot 10^{-1}$ ). For both cases the inaccuracy increases for larger fine-scale permeability contrasts. Off-diagonal terms show a small numerical error, as they deviate from zero.

Figure 4.16 presents the pressure fields, which results from loading by linear essential boundary conditions on a horizontal layered field composed out of  $4 \times 4$  periodic cells.

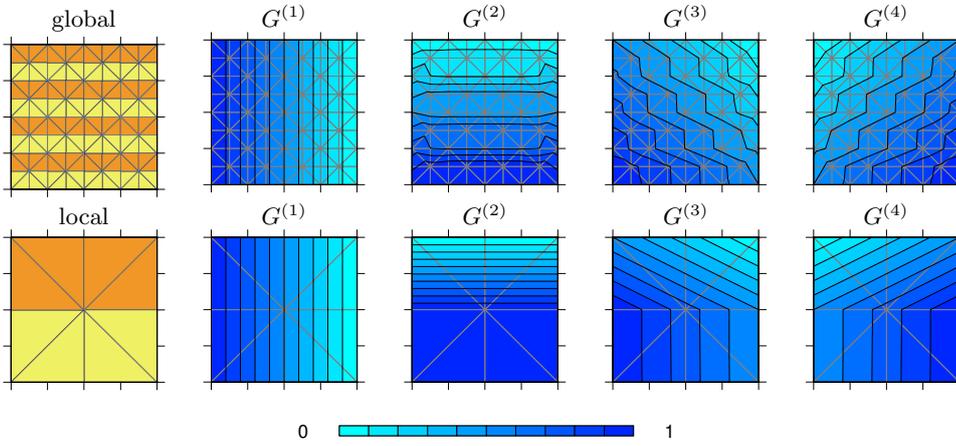


Figure 4.16: Global-local boundary conditions on a horizontally layered cell.

The upper pictures of this figure show the dissipation of the boundary effects. Still pressure-dissipation averaging obtains inaccurate results ( $k_{xx} = 5.05 \cdot 10^{-1}$ ,  $k_{xy} = 1.64 \cdot 10^{-16}$ ,  $k_{yx} = 3.64 \cdot 10^{-16}$ , and  $k_{yy} = 1.11 \cdot 10^{-1}$ ). The solution of the global problem could however be used to generate boundary conditions for a local problem. The pressure fields for this global-local procedure are given in the lower plots of figure 4.16. Pressure-dissipation averaging now computes accurate permeability tensor components ( $k_{xx} = 5.05 \cdot 10^{-1}$ ,  $k_{xy} = 2.55 \cdot 10^{-3}$ ,  $k_{yx} = 2.55 \cdot 10^{-3}$ , and  $k_{yy} = 1.99 \cdot 10^{-2}$ ). For this case global pressure results could directly be imposed on the local problem, because the mesh size of the local problem equals the mesh size of the global problem. The boundary conditions which were extracted from the global computation on this horizontal layered cell, generate a local flow field that nearly corresponds to the flow field for periodic boundary conditions on the periodic cell. For the horizontally layered cell these boundary conditions match oscillating boundary conditions, which follow from closure by a dimensionally reduced flow problem.

## 4.2 Multiscale formulation

The proposed multiscale formulation computes basis functions by solving local flow problems over patches of simplex elements. Only if the patch aligns with the small scale structure then the periodic boundary conditions match with the solution of dimensionally reduced flow computations. Single simplex elements do not meet this condition. This restriction becomes more severe in a hierarchical setting, where the averaging domain might only cover a small number of fine scale elements. The original two-level formulation for the multiscale basis functions, which relates the pressure on each level to the fine scale pressure, is used to interpolate the coarse scale results and to obtain estimates for the fine scale fluxes perpendicular to the element boundaries, which are used for the refinement criterion.

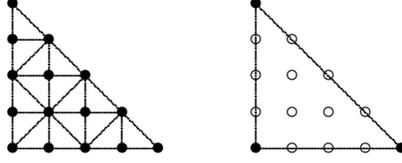


Figure 4.17: Direct multiscale basis function formulation.

Section 4.2.1 presents the formulation of the multiscale basis functions. Section 4.2.2 outlines the construction of the coarse grid operator.

### 4.2.1 Multiscale basis functions

Figure 4.17 illustrates the coarsening procedure according to the two-level multiscale finite element approach. The expression for the multiscale basis function reads

$$N_a^c = P_{ca} N_c^f, \quad (4.38)$$

where  $N_a^c$  denotes the coarse scale basis function attached to node  $a$ , and  $N_c^f$  expresses the fine scale basis function attached to node  $c$ . In multiscale finite element literature, the matrix coefficients  $P_{ca}$  are called basis function weights. In the previous sections the coefficients had constant values and where part of the interpolation matrix. Here their values will depend on the solution of local problems.

#### Variational formulation

Despite consistency of both formulations given by equations (4.8), (4.11) and equations (4.14), (4.16) for the one-dimensional test problem in the previous section, it can be shown that the consistent coarse-grid approach is not accurate, when a volume average of the permeability is adopted. The accuracy of the variational fine-grid formulation can be improved by using multiscale basis functions, where the coefficients of the interpolation matrix follow from a local flow analysis, and therefore better account for permeability variations. For a line element depicted in figure 4.18 the multiscale basis functions read

$$N_1^{k-1} = N_1^k + P_{31} N_3^k, \quad N_2^{k-1} = N_2^k + P_{32} N_3^k, \quad (4.39)$$

where the weights  $P_{31}^k$  and  $P_{32}^k$  need to be calculated.

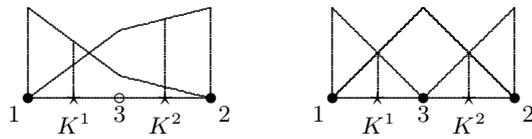


Figure 4.18: Multiscale basis functions for line elements.

Disregarding gravity on the test problem illustrated by figure 4.3, the continuity of flux over the fine-scale node gives  $K^1(p_1 - p_3) = K^2(p_3 - p_2)$ , where  $K^1$  is the permeability of the left element and  $K^2$  is the permeability of the right element. Pressures are given as nodal values. From this expression the pressure in the fine-scale node reads

$$p_3 = \frac{K^1}{K^1 + K^2} p_1 + \frac{K^2}{K^1 + K^2} p_2. \quad (4.40)$$

The multiscale basis functions that ensure continuity of flow can now be written as

$$N_1^{k-1} = N_1^k + \frac{K^1}{K^1 + K^2} N_3^k, \quad N_2^{k-1} = N_2^k + \frac{K^2}{K^1 + K^2} N_3^k, \quad (4.41)$$

and the interpolation matrix reads

$$[P_{ab}^k] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{K^1}{K^1 + K^2} & \frac{K^2}{K^1 + K^2} \end{bmatrix}. \quad (4.42)$$

Figure 4.18 shows the modified basis functions graphically. Using this interpolation matrix, variational coarsening now obtains an accurate coarse scale stiffness matrix from the fine scale stiffness matrix. The fine-scale matrix coefficients read

$$L_{ab}^k = \bigwedge_{e=1}^{n_e^k} \int_{\Omega^e} \frac{\rho k_r K}{\mu} \frac{\partial N_a^k}{\partial x} \frac{\partial N_b^k}{\partial x} d\Omega, \quad (4.43)$$

or

$$L_{ab}^k \approx \bigwedge_{e=1}^{n_e^k} w_m^e \left( \frac{\rho k_r K}{\mu} \frac{\partial N_a^k}{\partial \xi} \frac{\partial \xi}{\partial x} \frac{\partial N_b^k}{\partial \xi} \frac{\partial \xi}{\partial x} |J| \right)_m^e. \quad (4.44)$$

For elements with length 0.5 the fine-scale stiffness matrix can be written as

$$[L^k] = \frac{2\rho k_r}{\mu} \begin{bmatrix} K^1 & 0 & -K^1 \\ 0 & K^2 & -K^2 \\ -K^1 & -K^2 & K^1 + K^2 \end{bmatrix}. \quad (4.45)$$

Variational coarsening generates the coarse-scale matrix, which can be expressed as

$$L_{ab}^{k-1} = \bigwedge_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho k_r K}{\mu} P_{ca}^k \frac{\partial N_c^k}{\partial x} P_{db}^k \frac{\partial N_d^k}{\partial x} d\Omega, \quad (4.46)$$

approximated by

$$L_{ab}^{k-1} \approx \bigwedge_{e=1}^{n_e^k} P_{ca}^k w_m^e \left( \frac{\rho k_r K}{\mu} \frac{\partial N_c^k}{\partial \xi} \frac{\partial \xi}{\partial x} \frac{\partial N_d^k}{\partial \xi} \frac{\partial \xi}{\partial x} |J| \right)_m^e P_{db}^k. \quad (4.47)$$

Inserting expression (4.42) for the interpolation matrix gives

$$[L^{k-1}] = \frac{\rho k_r}{\mu} \frac{2}{K^1 + K^2} \begin{bmatrix} K^1 K^2 & -K^1 K^2 \\ -K^1 K^2 & K^1 K^2 \end{bmatrix}. \quad (4.48)$$

The right hand side of the discrete fine scale flow equation is written as

$$f_a^k = \sum_{e=1}^{n_e^k} \int_{\Omega^e} \frac{\rho k_r K}{\mu} \frac{\partial N_a^k}{\partial x} \rho g d\Omega, \quad (4.49)$$

approximated by

$$f_a^k \approx \sum_{e=1}^{n_e^k} w_m^e \left( \frac{\rho k_r K}{\mu} \frac{\partial N_a^k}{\partial \xi} \frac{\partial \xi}{\partial x} \rho g |J| \right)_m^e. \quad (4.50)$$

Inserting the fine-scale basis functions gives the fine-scale vector expressed by

$$[f^k] = \rho g \frac{\rho k_r}{\mu} \begin{bmatrix} -K^1 \\ K^2 \\ K^1 - K^2 \end{bmatrix}. \quad (4.51)$$

Variational coarsening produces the coarse-grid vector as

$$f_a^k = \sum_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho k_r K}{\mu} P_{ba}^k \frac{\partial N_b^k}{\partial x} \rho g d\Omega, \quad (4.52)$$

or

$$f_a^k \approx \sum_{e=1}^{n_e^k} P_{ba}^k w_m^e \left( \frac{\rho k_r K}{\mu} \frac{\partial N_b^k}{\partial \xi} \frac{\partial \xi}{\partial x} \rho g |J| \right)_m^e. \quad (4.53)$$

Inserting the fine-scale basis functions and the expression for the interpolation matrix, according to equation (4.42), gives

$$[f^{k-1}] = \frac{\rho k_r}{\mu} \frac{2\rho g}{K^1 + K^2} \begin{bmatrix} -K^1 K^2 \\ K^1 K^2 \end{bmatrix}. \quad (4.54)$$

This result is also obtained from a homogeneous coarse-scale formulation and an effective harmonic mean permeability  $\langle K \rangle = 2(K^1 K^2)/(K^1 + K^2)$  sampled in a coarse-scale integration point. For a harmonic average, low conductivities dominate the coarse-grid average. This example indicates the importance of the way to take conductivity extremes into account.

In conclusion, the multiscale formulation, with interpolation weights given by the prolongation coefficients of equation (4.42), captures fine-scale behavior of the solution on a coarse mesh. Solving the coarse-scale problem will give the exact flow rates over the coarse-scale elements. The coarse-scale nodal values of the pressure are obtained with fine-scale accuracy. The fine-scale variation of the pressure is found by interpolation of the coarse-scale nodal values with the interpolation operator given by equation (4.42). This conclusion only holds for the one-dimensional case as will be explained next.

The proposed multiscale basis functions on a quadrilateral patch of triangular elements, follow from bi-linear boundary conditions. For a unit square domain these conditions coincide with the definition of the quadrilateral basis functions and read

$$B_1 = (1-x) \cdot (1-y), \quad B_2 = x \cdot (1-y), \quad B_3 = x \cdot y, \quad B_4 = (1-x) \cdot y. \quad (4.55)$$

Figure 4.19 shows multiscale basis functions on a patch of eight elements generated

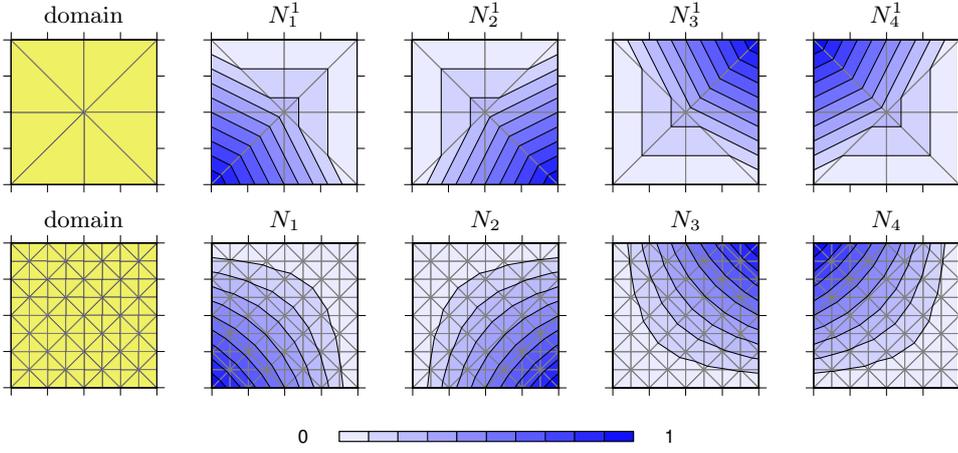


Figure 4.19: Multiscale basis functions on a patch of two triangular parent elements.

out of two triangular parent elements. Standard coarse scale basis functions over the patch of quadrilateral elements follow from the fine scale basis functions as

$$\begin{aligned}
 N_1^{k-1} &= N_1^k + \frac{1}{2}N_6^k + \frac{1}{2}N_9^k + \frac{1}{4}N_5^k, \\
 N_2^{k-1} &= N_2^k + \frac{1}{2}N_6^k + \frac{1}{2}N_7^k + \frac{1}{4}N_5^k, \\
 N_3^{k-1} &= N_3^k + \frac{1}{2}N_7^k + \frac{1}{2}N_8^k + \frac{1}{4}N_5^k, \\
 N_4^{k-1} &= N_4^k + \frac{1}{2}N_8^k + \frac{1}{2}N_9^k + \frac{1}{4}N_5^k.
 \end{aligned} \tag{4.56}$$

This formulation restores symmetry on the patch. The span of basis functions  $N_1$  and  $N_3$  increases compared with the standard formulation, which was depicted by figure 4.8.

### Sampling domain

The values of the multiscale weights often strongly depend on the closure of the local problem and the discretization size of this problem. These effects become more dominant when the averaging volume is small compared to the scale of parameter variation. Oversampling reduces the boundary effects by posing the boundary conditions on a sampling domain, which extends the patch of simplex elements as depicted by figure 4.20. The overlapping domains are closed by either linear or oscillating essential boundary conditions. Multiscale basis functions, which apply to the patch, now follow from the oversampled multiscale basis functions as

$$N_a = c_{ab}N_b^*, \tag{4.57}$$

where  $N_b^*$  denote the base functions over the sampling domain,  $N_a$  is the base function over the element patch and  $c_{ab}$  are the coefficients, which follow from  $N_a(x_b) = \delta_{ab}$ .

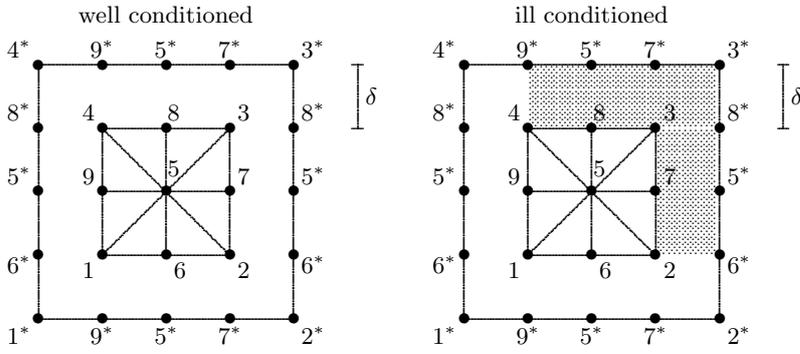


Figure 4.20: Oversampling domain.

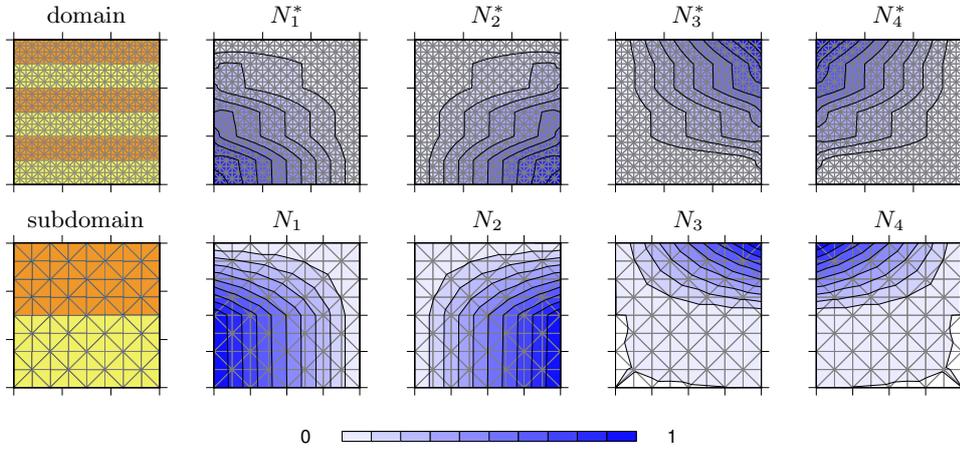


Figure 4.21: Multiscale basis functions on a horizontally layered field.

As an example

$$\begin{bmatrix} N_1^*(x_1) & N_2^*(x_1) & N_3^*(x_1) & N_4^*(x_1) \\ N_1^*(x_2) & N_2^*(x_2) & N_3^*(x_2) & N_4^*(x_2) \\ N_1^*(x_3) & N_2^*(x_3) & N_3^*(x_3) & N_4^*(x_3) \\ N_1^*(x_4) & N_2^*(x_4) & N_3^*(x_4) & N_4^*(x_4) \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{12} \\ c_{13} \\ c_{14} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (4.58)$$

gives  $N_1 = c_{11}N_1^* + c_{12}N_2^* + c_{13}N_3^* + c_{14}N_4^*$ .

Figure 4.21 presents multiscale basis functions, which follow from linear closure of a horizontally layered domain. This domain over samples the subdomain that captures a single periodic cell. The bottom row presents the extracted multiscale basis functions on the reduced domain. Results are given for a permeability contrast of  $10^2$ .

Reduction of basis functions from a domain to a subdomain might pose an ill-conditioned set of equations. The picture right in figure 4.20 shows an example for this. Here the coarse scale node 3, is isolated from the subdomain by a low permeable

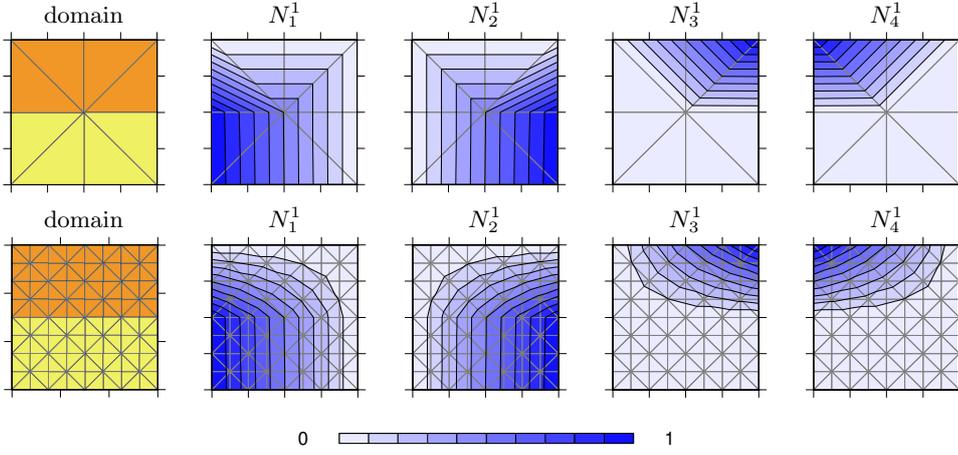


Figure 4.22: Multiscale basis functions for a horizontal layered cell.

zone. For this situation the components of the third column of the matrix given by equation (4.58) become small and  $c_{13}$  is undetermined. For this reason a second approach will be considered. This approach collapses the mesh over the boundary zone and generates dimensionally reduced flow problems. Figure 4.20 also illustrates the second approach that generates conforming basis functions by first computing oscillating boundary conditions. The algebraic form of the reduction problem associated with the bottom boundary of the patch can be written as

$$\begin{bmatrix} a_{11} & & a_{15} & a_{16} & & a_{19} \\ & a_{22} & a_{25} & a_{26} & a_{27} & \\ a_{51} & a_{52} & a_{55} & a_{56} & a_{57} & a_{59} \\ a_{61} & a_{62} & a_{65} & a_{66} & & \\ & a_{72} & a_{75} & & a_{77} & \\ a_{91} & & a_{95} & & & a_{99} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_5 \\ p_6 \\ p_7 \\ p_9 \end{bmatrix} = \begin{bmatrix} Q_1 \\ Q_2 \\ Q_5 \\ Q_6 \\ Q_7 \\ Q_9 \end{bmatrix}. \quad (4.59)$$

The lumped expression for the pressure at node 6 then follows from

$$(a_{61} + a_{51}) p_1 + (a_{62} + a_{52}) p_2 + (a_{65} + a_{66} + a_{55} + a_{56}) p_6 = 0. \quad (4.60)$$

The lumping process is illustrated by figure C.8. The connectivities to nodes on the opposite side of the boundary generalize this expression as

$$(a_{61} + a_{51} + a_{5^*1}) p_1 + (a_{62} + a_{52} + a_{5^*2}) p_2 + (a_{65} + a_{65^*} + a_{66} + a_{55} + a_{5^*5} + a_{56} + a_{5^*6}) p_6 = 0. \quad (4.61)$$

Figure 4.22 presents fully-resolved and over-resolved multiscale basis functions for a horizontally layered cell. These functions follow from a local flow computation with oscillating boundary conditions.

Appendix B presents multiscale basis functions for four periodic fields. The appendix includes results for the horizontally layered field (figure B.12), the checkerboard

field (figure B.13), the domain with inclusions (figure B.14), and the diagonal layered field (figure B.15). The permeability contrast for all domains equals  $10^2$  and results are obtained for an over-resolved finite element discretization.

### 4.2.2 Coarse mesh equations

Oversampling obtains best results for the computation of equivalent permeability. However the multiscale basis functions are no longer continuous over their domain boundaries and the span increases, if these basis functions are applied to obtain the coarse grid operator. The method sometimes even produces basis function weights significantly larger than one on interpolation points for ill-conditioned situations. Sampling directly over simplex element enforces the compatibility conditions. Using a patch of elements already can be seen as oversampling, as the node in the center of the patch is connected to all coarse grid nodes. Pressure-dissipation averaging prevents new connectivities to occur. Alternatively a dual mesh can be constructed. The stiffness matrix then follows as the sum of the algebraic equations derived on the original mesh and the dual mesh. This stiffness matrix also includes the extra connectivities. The dual mesh corrects mesh dependency and the sum of both stiffness matrices is able to incorporate non-diagonal terms in the permeability tensor at all nodes. The costs of this procedure is relatively low as the dual meshes need not be constructed on the finest level.

#### Pressure-dissipation formulation

Variational coarsening generates the coarse-scale matrix, given by

$$L_{ab}^{k-1} = \bigwedge_{e=1}^{n_e^{k-1}} \int_{\Omega^e} \frac{\rho k_r K_{ij}}{\mu} P_{ca}^k \frac{\partial N_c^k}{\partial x_i} P_{db}^k \frac{\partial N_d^k}{\partial x_j} d\Omega, \quad (4.62)$$

approximated by

$$L_{ab}^{k-1} \approx \bigwedge_{e=1}^{n_e^k} w_m^e P_{ca}^k \left( \frac{\rho k_r K_{ij}}{\mu} \frac{\partial N_c^k}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_d^k}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} |J| \right)_m P_{db}^k. \quad (4.63)$$

The right hand side of the discrete flow equation is written as

$$f_a^{k-1} = \bigwedge_{e=1}^{n_e^k} \int_{\Omega^e} \frac{\rho k_r K_{ij}}{\mu} P_{ca}^k \frac{\partial N_c^k}{\partial x_i} \rho g_j d\Omega, \quad (4.64)$$

or

$$f_a^{k-1} \approx \bigwedge_{e=1}^{n_e^k} w_m^e P_{ca}^k \left( \frac{\rho k_r K_{ij}}{\mu} \frac{\partial N_c^k}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \rho g_j |J| \right)_m. \quad (4.65)$$

Evaluating the coarse scale operator over a subdomain of fine scale elements is computationally expensive. Pressure-dissipation averaging obtains the same coarse grid operator if the procedure applies to all fine scale elements, which form the coarse scale element. Multiscale basis functions modify expression (4.35) as

$$\phi_{ij}^{(ab)} = \bigwedge_{e=1}^{n_e^{k-1}} w_m^e \left( P_{ca} p_a^{(a)k} \frac{\rho k_r}{\mu} \frac{\partial N_c^k}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i} \frac{\partial N_d^k}{\partial \xi_l} \frac{\partial \xi_l}{\partial x_j} P_{db} p_b^{(b)k} |J| \right)_m. \quad (4.66)$$

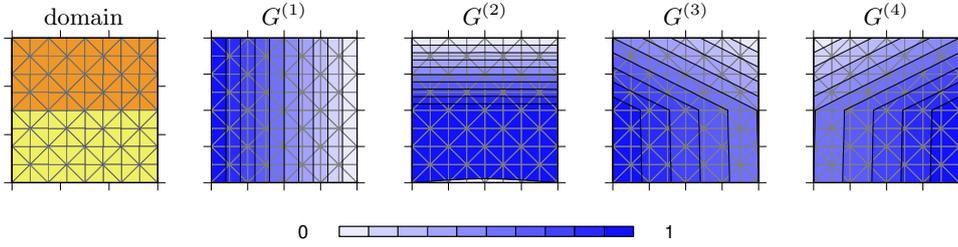


Figure 4.23: Oversampled solution on a horizontally layered field.

The linear loading case for the patch given by figure 4.7, can be constructed out of coarse scale basis functions. These conditions reads

$$G^{(1)} = N_1^{k-1} + N_4^{k-1}, \quad G^{(2)} = N_1^{k-1} + N_2^{k-1}, \quad (4.67)$$

or alternatively

$$G^{(3)} = N_1^{k-1} + \frac{1}{2}N_2^{k-1} + \frac{1}{2}N_4^{k-1}, \quad G^{(4)} = N_2^{k-1} + \frac{1}{2}N_1^{k-1} + \frac{1}{2}N_3^{k-1}. \quad (4.68)$$

Figure 4.23 shows the reconstruction of the loading cases by adding the multiscale basis functions for the horizontally layered periodic cell. Pressure-dissipation averaging extracts accurate equivalent permeability tensor component for this case ( $k_{xx} = 5.05 \cdot 10^{-1}$ ,  $k_{xy} = -8.31 \cdot 10^{-16}$ ,  $k_{yx} = -1.76 \cdot 10^{-15}$ , and  $k_{yy} = 1.98 \cdot 10^{-2}$ ).

Appendix A includes the reconstruction of the loading cases for the horizontally layered field (figure A.2), the checkerboard field (figure A.3), the domain with inclusions (figure A.4), and the diagonally layered field (figure A.5). Basis functions were obtained on overlapping (oversampled) subdomains and non-overlapping (oscillating) subdomains. All results are compared with periodic loading of the periodic cell in accompanying pictures. The appendix also presents the results for the equivalent permeability in table A.1, table A.2, table A.3, and table A.4. These tables show that a fully resolved horizontally layered cell and a fully resolved domain with a single inclusion already provide accurate equivalent permeability tensor components. The checkerboard field requires an over-resolved finite element mesh, which captures the singular points. A fine mesh captures the structure of the diagonally layered field more accurately, and improves the prediction of the equivalent permeability components. Measuring the difference of equivalent permeability predictions provides a second refinement criterion

$$|\text{tr}(\tilde{K})^{\text{ov}} - \text{tr}(\tilde{K})^{\text{os}}| > c_2 \cdot \text{tr}(\tilde{K})^{\text{os}}, \quad (4.69)$$

where  $\text{tr}(\tilde{K})^{\text{ov}}$  denotes the trace of the equivalent permeability found by oversampling the subdomain. The first invariant  $\text{tr}(\tilde{K})^{\text{os}}$  follows from imposing oscillating boundary conditions. The trace of the tensor matrix sums the diagonal components of the matrix.

### 4.3 Multigrid formulation

Domain decomposition methods like multigrid algorithms can be used as complete iterative algorithms to solve sets of equations directly or may be used as preconditioners

for Krylov subspace methods like the conjugate gradient method. Domain decomposition algorithms can be divided in two classes. Overlapping domain algorithms refer to Schwarz methods and include multigrid methods. Non-overlapping domain algorithms refer to Schur complement methods and provide substructuring techniques. Additive Schwarz methods solve the system of algebraic equations on overlapping subdomains and do not use updated values on the subdomain boundaries, which follow from computations on neighboring subdomains. These methods are suitable for parallel implementation and relate to Jacobi solvers. Multiplicative Schwarz methods use updated boundary information as soon as it becomes available. These methods relate to Gauss-Seidel algorithms and need coloring for parallel computation. The convergence rate for the additive Schwarz method is lower than the rate for the multiplicative Schwarz method. In general convergence is poor if there is no overlap domain ( $\delta = 0$ ) in figure 4.20. Multiscale finite element methods can be classified as non-iterative domain decomposition methods. In classical multigrid, local iterative solvers like Jacobi or Gauss-Seidel remove the high frequency components of the error. Multilevel overlapping domain decomposition applies Schwarz smoothers. Jacobi multigrid follows as a multilevel additive Schwarz method on matching grids and subdomains with minimal overlap as depicted in figure 4.24. This figure shows two overlapping subdomains that consist of the support of the basis function associated with nodes 6 and 7. The overlap consists of 2 elements (3,6,7) and (6,7,11). Multiplicative Schwarz domain decomposition gives Gauss-Seidel multigrid on domains with minimal overlap.

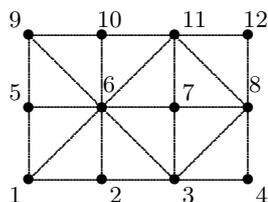


Figure 4.24: Minimal overlapping subdomains.

Appendix C presents the linear multigrid algorithm and generalizes this algorithm to the non-linear Full Approximation Storage (FAS) algorithm. The multigrid method presented in this section extends the two-level multiscale finite element method to a multilevel algorithm. The proposed geometric multigrid method is a physically based technique needed in an adaptive formulation. It operates on a predefined grid hierarchy like the one shown by Figure 4.1. The section considers the matrix properties of the fine scale operator. The construction of the coarse scale operator and intergrid transfer operators will be outlined and procedures for measuring the convergence behavior of the algorithm will be given.

### Fine grid operator

Multigrid methods are well established for solving self-adjoint boundary value problems on fields of coefficients that show mild variations over the domain. The self-adjoint operator for these problems corresponds to its adjoint, or equivalently the operator

matrix is a Hermitian matrix. This matrix is its own conjugate transpose, and if the Hermitian matrix is real then the matrix is symmetric. Matrices of real discretized self-adjoint boundary value problems are symmetric, sparse, and weakly diagonal dominant. For weakly diagonal dominant matrices the diagonal element is at least as large as the sum of the absolute values of the off-diagonal elements. It can be shown that if a matrix is symmetric and weakly diagonal dominant then the matrix is positive definite. For a positive definite matrix the following conditions apply: for all  $u \neq 0$  the relation  $u^T A u > 0$  holds and  $A = A^T$ . An M-matrix forms a special class, this matrix is symmetric and positive definite with non-positive off-diagonal terms.

The fine scale operator matrix of the flow problem can be expressed by a finite difference stencil. The finite difference formulation follows from a uniform finite element mesh. Two configurations of the center nodal have to be investigated. Figure 4.25 shows these configurations. Components of the conductivity matrix which follow from

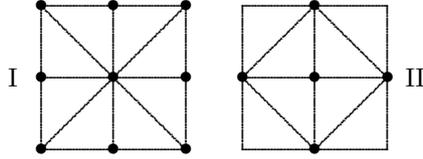


Figure 4.25: Uniform triangular mesh configurations.

the finite element discretization of the diffusive term read

$$S_{ab} = \sum_{e=1}^{n_e} \int_{\Omega^e} \frac{\partial N_a}{\partial x_i} K_{ij} \frac{\partial N_b}{\partial x_j} d\Omega \quad (4.70)$$

For configuration I the finite difference stencil reads

$$\begin{aligned} ]S_h^I [ = K_{xx} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} + K_{yy} \begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix} \\ + \frac{K_{xy}}{2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} + \frac{K_{yx}}{2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}. \quad (4.71) \end{aligned}$$

In two dimensions the conductivity matrix is independent of the size of the uniform element distribution. This stencil captures the off-diagonal terms in the permeability matrix. However for boundary edge nodes configuration II applies. The finite difference stencil for these nodes reads

$$]S_h^{II} [ = K_{xx} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} + K_{yy} \begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix}. \quad (4.72)$$

Because of the absence of the diagonal connectivities this stencil does not capture the effects of off-diagonal permeability terms. Configuration II produces an M-matrix,

configuration I does not produce such a matrix if off-diagonal tensor components are considered.

Variable coefficient problems and variable mesh size problems generate anisotropic algebraic problems, and weak connections in one direction. Local smoothers reduce the error in the direction of the strong connection but do not reduce the high frequency error in the direction of the weak connection. As a result multigrid solvers show poor performance as convergence factors degrade. Semicoarsening, block or line relaxation improves the convergence rate of the multigrid solvers on anisotropic problems. Semicoarsening applies point relaxation and coarsens the mesh in the direction of the high connectivity because multigrid convergence can only be expected in this direction. Semicoarsening does not use weakly connected neighbors for intergrid transfer operations. Line or block relaxation solves the strongly coupled unknowns directly and applies sequential standard coarsening. Variable-coefficient problems require more sophisticated multigrid methods if the amplitude of coefficient varies strongly over the field and the frequency variation is high.

### Intergrid transfer operator

Multiscale basis functions provide intergrid transfer operators. The intergrid transfer operators are based on a hierarchical formulation of the multiscale basis functions. Figure 4.26 illustrates the sequential procedure. The sequential expression for the

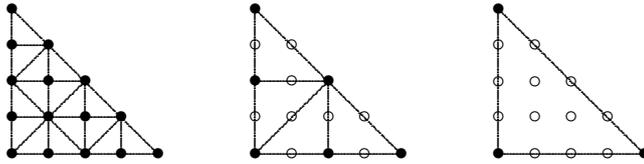


Figure 4.26: Sequential multiscale basis functions formulation.

multiscale basis function reads

$$N_a^{k-1} = P_{ca}^k N_c^k, \quad N_a^{k-2} = P_{ca}^{k-1} N_c^{k-1}, \quad (4.73)$$

where  $k$  indicates the mesh level. Prolongation and restriction operators transfer information over the hierarchy of grids. The prolongation operator  $[P_{k-1}^k]$ , which interpolates coarse scale results to the fine scale, includes the components of the intergrid transfer operator. The restriction matrix  $[R_k^{k-1}]$  injects the fine scale information to the coarse scale. The matrix follows as the (weighted) transpose of the prolongation matrix, as a variational formulation will be applied.

### Coarse grid operator

The multiscale homogenization formulation provides equivalent coarse-scale behavior. Coarse grid operators follow from pressure-dissipation homogenization on non-oversampled subdomains, and multiscale basis function loading conditions. Coarse-grid discretization produces coarse-grid operators. Hierarchical scaling, implicit in algebraic

multigrid and explicit in geometric multigrid, produces less accurate equivalent permeability tensors due to the sequential closure of local flow problems. More accurate results are obtained if the fine scale information is brought to each scale. The method approximates variational coarsening, and provides an alternative for more complex problems. The coarse-grid operator and coarse-grid force vector then follow from the fine-scale operator and intergrid transfer operators as

$$[L^{k-1}] = [R_k^{k-1}] [L^k] [P_{k-1}^k], \quad [f^{k-1}] = [R_k^{k-1}] [f^k]. \quad (4.74)$$

This expression is known as the Galerkin approximation. Variational coarsening uses the adjoint of the prolongation operator as a restriction operator

$$[R_k^{k-1}] = [P_{k-1}^k]^T. \quad (4.75)$$

In the Galerkin formulation the restriction matrix is not weighted by the sum of its rows.

### Full approximation storage algorithm

Non-linear algebraic equations can be solved by the full approximation storage algorithm (FAS). Here a non-linear Gauss-Seidel relaxation is mostly used to smooth the non-linear algebraic equations. FAS modifies the coarse scale equations with a  $\tau$ -correction (appendix C). This correction improves the accuracy of the coarse scale solution. If the system matrix is a linear operator, then the FAS scheme reduces to the linear correction scheme. Adaptive methods use estimates of the discretization error as a criterion for mesh refinement. In a multilevel setting the discretization errors can be estimated on several levels and the error on the finest level can be extrapolation. The FAS scheme provides the approximation on each level, which can be used directly. The proposed multigrid algorithm is based on the FAS scheme and applies Gauss-Seidel iterations to smooth the high frequency modes of the error. A V-cycle schedule performs the multigrid iterations. The algorithm applies non-weighted restriction of the defect, weighted restriction of the primary variable and multiscale basis function interpolation of the unknown pressure.

Multigrid is an iterative method and needs a stopping criterion. This criterion will be based on the algebraic error. The final algebraic error has to be below the level of discretization error. The algebraic error is the difference between the solution of the algebraic equation and the solution of the exact discrete solution. The discretization error compares the solution of the algebraic equations to the solution of the continuous problem given by differential equations in discrete points. The maximum or infinity norm and Eulidean or 2-norm can be used to measure the error of the solution. Here the discrete  $L^2$  norm of a function will be used. This Eulerian norm reads

$$\|e\|_2 = \sqrt{\sum_{e=1}^{n_e} \int_{\Omega_e} (N_a \tilde{u}_a - N_a u_a)^2 d\Omega}, \quad (4.76)$$

where  $\tilde{u}_a$  denotes a fine scale reference solution. A discrete error norm which shows a decrease of a factor four, on a uniform two-dimensional mesh and standard coarsening,

Table 4.1: Error norm checkerboard problem.

cycle	$\ e\ _2^1$	$\ e\ _2^2$	$\ e\ _2^3$	$\ e\ _2^4$
1	$2.03 \cdot 10^{+0}$	$1.06 \cdot 10^{+0}$	$7.20 \cdot 10^{-1}$	$6.10 \cdot 10^{-1}$
2	$2.03 \cdot 10^{+0}$	$1.05 \cdot 10^{+0}$	$4.37 \cdot 10^{-1}$	$4.41 \cdot 10^{-2}$
3	$2.03 \cdot 10^{+0}$	$1.05 \cdot 10^{+0}$	$4.29 \cdot 10^{-1}$	$3.31 \cdot 10^{-3}$
4	$2.03 \cdot 10^{+0}$	$1.05 \cdot 10^{+0}$	$4.28 \cdot 10^{-1}$	$2.49 \cdot 10^{-4}$
5	$2.03 \cdot 10^{+0}$	$1.05 \cdot 10^{+0}$	$4.28 \cdot 10^{-1}$	$1.87 \cdot 10^{-5}$
6	$2.03 \cdot 10^{+0}$	$1.05 \cdot 10^{+0}$	$4.28 \cdot 10^{-1}$	$1.40 \cdot 10^{-6}$
7	$2.03 \cdot 10^{+0}$	$1.05 \cdot 10^{+0}$	$4.28 \cdot 10^{-1}$	$1.05 \cdot 10^{-7}$
8	$2.03 \cdot 10^{+0}$	$1.05 \cdot 10^{+0}$	$4.28 \cdot 10^{-1}$	$7.34 \cdot 10^{-9}$

Table 4.2: Residual norm checkerboard problem.

cycle	$\ r\ _2^1$	$\ r\ _2^2$	$\ r\ _2^3$	$\ r\ _2^4$
1	$8.19 \cdot 10^{-07}$	$1.76 \cdot 10^{-03}$	$9.94 \cdot 10^{-03}$	$1.14 \cdot 10^{-02}$
2	$1.02 \cdot 10^{-12}$	$4.72 \cdot 10^{-06}$	$3.26 \cdot 10^{-04}$	$8.74 \cdot 10^{-04}$
3	$1.90 \cdot 10^{-16}$	$1.31 \cdot 10^{-08}$	$1.09 \cdot 10^{-05}$	$6.64 \cdot 10^{-05}$
4	$5.25 \cdot 10^{-16}$	$3.71 \cdot 10^{-11}$	$3.67 \cdot 10^{-07}$	$5.00 \cdot 10^{-06}$
5	$5.25 \cdot 10^{-16}$	$1.07 \cdot 10^{-13}$	$1.24 \cdot 10^{-08}$	$3.76 \cdot 10^{-07}$
6	$5.25 \cdot 10^{-16}$	$1.07 \cdot 10^{-13}$	$4.15 \cdot 10^{-10}$	$2.83 \cdot 10^{-08}$
7	$5.25 \cdot 10^{-16}$	$1.07 \cdot 10^{-13}$	$1.40 \cdot 10^{-11}$	$2.13 \cdot 10^{-09}$
8	$5.25 \cdot 10^{-16}$	$1.07 \cdot 10^{-13}$	$1.40 \cdot 10^{-11}$	$1.60 \cdot 10^{-10}$

as the resolution doubles indicates second order discretization. The residual norm is given by

$$\|r\|_2 = \sqrt{\sum_{e=1}^{n_e} \int_{\Omega_e} (N_a r_a)^2 d\Omega}. \quad (4.77)$$

If the ratio of the residual norm  $\|r^{r+1}\|_2 / \|r^r\|_2$  approaches a nearly constant ratio after a number of cycles then this ratio estimates the asymptotic convergence factor of the algorithm. A sharp increase at the end of the cyclic iterations indicates that the algebraic approximation is accurate to near machine precision. The smoothing rate of standard relaxation schemes indicates the convergence factor for reducing the oscillatory modes of the discrete error. The overall convergence factor for a multigrid scheme should be small and independent of the discretization size. For homogeneous problems convergence factors of 0.05 can be obtained, for variable coefficient problems convergence coefficients in the 0.2-0.3 range can be expected. Table 4.1 presents the error norm per cycle for the checkerboard problem already shown by figure 4.12.

Results were obtained for a hierarchy of finite element meshes consisting of 2, 8, 32, 128, 512 and 2048 triangular elements. The error norm is presented for computations with meshes that consist of 25, 81, 289 and 1089 nodes at the finest level. Gauss-Seidel iterations smooth the error on each level by two pre-smoothing and two post-smoothing operations. All results are compared with the final result of the 1089 node computation. The error norm reduces linearly for fixed refinement. Table 4.2 presents the residual norm for the checkerboard problem per cycle. The convergence factor is 0.03 for the third computation and equals 0.07 for the fourth analysis.

# Chapter 5

## Model verification

For partly saturated porous media the permeability depends on the saturation of the medium by the fluid phase. The saturation itself depends on the pore-water pressure [71]. A constitutive model captures this behavior and introduces non-linearities into the flow equation. Stable flow conditions occur, if the unsaturated zone is on top of the saturated zone, as in phreatic surface problems. Depending on the constitutive equations, infiltration problems may cause instable flow conditions. Small perturbations in the permeability field or porosity field for instance may trigger these instabilities and the infiltration fronts then show fingering. Transport of a solute is coupled to the flow of the fluid for which the solute contributes as a single component. The fluid velocity is present in the convective term and scales the hydrodynamic dispersivity in the diffusive term of the transport equation. Fully coupled flow and solute transport equations capture density driven flow. Here, the fluid density, present in the flow equation, depends on the solute mass fraction. For density driven flow stable flow situations occur if the fluid density decreases with elevation height. Depending on the hydrodynamic dispersion, instable situation may prevail if a denser fluid overlays a less dense fluid. For this situation convective cells may develop and even for the case without any diffusion, mixing takes place by convection. Heat transport gives rise to fully coupled flow and transport equations, as the fluid density and fluid viscosity both depend on pressure and temperature. Unstable flow conditions may occur in a porous medium that is heated from below.

This chapter considers the implementation of the saturated and unsaturated flow model, coupled with solute transport and heat transport. The accuracy of the proposed adaptive refinement formulation will be verified for a number of well known benchmark problems by comparing its results with fine grid results. The efficiency of the algorithm follows from the number of degrees of freedom that provide the level of accuracy. For this the  $L_2$  Euclidean norm will be used throughout the section. Section 5.1 verifies the modeling of saturated and unsaturated flow under stable and instable conditions. Section 5.2 verifies the modeling of solute and heat transport for similar conditions.

## 5.1 Saturated and unsaturated flow

This section verifies the adaptive multilevel formulation for saturated and unsaturated flow. The constitutive equations are given by either an artificial model or the Van Genuchten material model. The problems that will be investigated, impose prescribed pressure boundary conditions, prescribed flux conditions and seepage conditions. The seepage conditions pose a prescribed pressure, if outflow occurs, and impose closed boundary conditions, if the pressure becomes negative under unsaturated conditions.

Section 5.1.1 presents the Lamb and Whitman problem [39, 51, 19]. The problem involves stable saturated flow through an earth dam, but is solved in this section in a partly saturated domain. The modified simulation focuses on the position of the phreatic line and length of the internal seepage face along the material interface. Section 5.1.2 presents Forsyth's problem. This problem involves infiltration in a large caisson filled with four different porous materials [40]. The numerical simulation focuses on the infiltration process and on the effect of material contrasts.

### 5.1.1 Lamb and Whitman's problem

Lamb and Whitman considered saturated flow through a cross section of an earth dam resting on an impervious foundation [51]. A rock toe in the dam prevents water from eroding the construction. For this reason, the toe is constructed out of a coarser material. The rock toe is more permeable than the base material of the dam, which reduces the water pressure distribution in the dam. Water pressures will have a large impact on the stability of the dam and need to be considered in the design of the construction. Lamb and Whitman constructed a flow net to simulate the steady state seepage through the dam. An iterative procedure had to be used because the flow domain is closed by a phreatic line, and the position of this line is not known in advance. The pressure head is zero along the phreatic line, so here the total head equals the elevation head.

Figure 5.1 shows the layout of Lamb and Whitman's earth dam problem. The

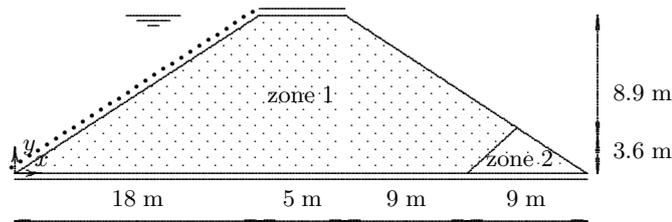


Figure 5.1: Lamb and Whitman's problem definition.

base of the dam is 41 m, its height is 12.5 m. The base material of the dam has a permeability of 13.13 m/d, the permeability of the rock toe is set to  $13.13 \cdot 10^8$  m/d. On the upstream side the water is at reservoir level of 12 m. At the downstream side, the water is at soil surface level. Here the total head equals zero. The simulation will follow the benchmark problem outlined in the Plaxflow manual [19].

The steady state situation is obtained by a transient computation. The simulation starts with a fully unsaturated cross section, and then computes the filling of the dam after a gradual rise of river water level. The numerical simulation follows the physical process and provides time dependent water pressure fields, which can be used in a decoupled stability assessment. The model is also able to compute water pressures in the unsaturated zone. In general, these negative pressures will stiffen the soil. The influence of flow through the unsaturated zone is reduced however by imposing a constitutive model which provides a sharp transition zone between the saturated and unsaturated domain. This makes the comparison to the original solution easier. The simulation time needed to reach the steady state condition is about 15 days. The hierarchy of finite element meshes consists of 30, 99, 357, 1353 nodes, and 48, 176, 672, 2624 elements. On the most coarse level the size of the elements ranges between 3.816 m up till 5.184 m in x-direction and 2.65 m till 3.6 m in y-direction. The simulation period of  $1.5 \cdot 10^6$  s was divided in 150 steps of  $1.0 \cdot 10^4$  s each.

Initial conditions specify a hydrostatic fully unsaturated field. These conditions are given by

$$p^0 = -\gamma \cdot y, \quad (5.1)$$

where  $\gamma$  equals  $10^4$  Pa. The pressure at the top of the dam then equals  $-1.25 \cdot 10^5$  Pa at the beginning of the simulation.

Pressure conditions apply to the inner and the outer slope of the dam, and simulate the reservoir and polder influence. Boundary conditions on the outer slope specify water loading, and read

$$\begin{cases} p = -\gamma \cdot y + 0.125 \cdot t & \text{if } 0 < t \leq t^1 \cap p \geq 0 \\ q_x n_x + q_y n_y = 0 & \text{for } 0 < t \leq t^1 \cap p < 0 \\ p = -\gamma \cdot y + \gamma h & \text{if } t > t^1 \end{cases} . \quad (5.2)$$

At  $t^1 = 10^6$  s the water level of the reservoir equals 12.5 m, which corresponds to the construction height  $h$  of the dam. The boundary condition at this outer slope generalizes the seepage condition. Outflow of groundwater only occurs for saturated conditions, on the unsaturated part of the boundary outflow of water is restricted. A prescribed pressure condition holds for the submerged part of the boundary. Boundary conditions on the inland slope specify either a no-flow condition or prescribe a pressure condition. No-flow conditions apply on unsaturated parts of the boundary and pressure conditions hold for parts of the boundary where outflow takes place. The position of the separated boundary parts is not known in advance and will be calculated automatically. These geometrically nonlinear boundary conditions read

$$\begin{cases} p = 0 & \text{if } q_x n_x + q_y n_y \geq 0 \\ q_x n_x + q_y n_y = 0 & \text{for } p < 0 \end{cases} . \quad (5.3)$$

At the bottom of the model no-flow conditions apply. The no-flow conditions are given by  $q_y = 0$ .

Table 5.1 lists the material parameters and their values. The transition from satu-

Table 5.1: Material parameters for Lamb and Whitman's problem.

parameter	value
permeability soil $\kappa_1$	$1.52 \cdot 10^{-11} \text{ m}^2$
permeability rock $\kappa_2$	$1.52 \cdot 10^{-7} \text{ m}^2$
porosity soil and rock $n_1, n_2$	0.35
density water $\rho$	$10^3 \text{ kg/m}^3$
viscosity water $\nu$	$10^{-3} \text{ Pas}$
compressibility skeleton $\alpha_p$	$0 \text{ m}^2/\text{N}$

rated to unsaturated condition is obtained by a linearized material model, given by

$$S = \begin{cases} S_s & \text{if } \psi > 0 \\ S_s + \frac{\psi}{|\psi^*|} \cdot (S_s - S_r) & \text{for } \psi^* \leq \psi \leq 0, \\ S_r & \text{if } \psi < \psi^* \end{cases}, \quad (5.4)$$

and

$$k_r = \begin{cases} 1 & \text{if } \psi > 0 \\ 10^{\frac{4\psi}{|\psi^*|}} & \text{for } \psi^* \leq \psi \leq 0, \\ 10^{-4} & \text{if } \psi < \psi^* \end{cases}, \quad (5.5)$$

where the pressure head reads  $\psi = p/\rho g$  and  $\psi^*$  denotes a minimum threshold value, which relates to the height of a coarse grid element. Simulations were carried out for a residual saturation  $S_r$  of 0.1, a maximum saturation  $S_s$  of 1.0, and a threshold value  $\psi^*$  of 0.5 m. The relative permeability varies over four orders of magnitude, and the permeability contrast over the rock and soil material equals  $10^8$ . The gravity constant is  $10.0 \text{ m/s}^2$ .

Figure 5.2 presents the pressure field for a number of time steps. The pictures only show the saturated part of the flow domain. Although a relatively sharp transition zone is enforced, groundwater flow occurs above the phreatic surface as well. This situation was not captured by Lamb and Whitman, but occurs in real physical systems. The pictures show the time dependent filling process of the dam. The position of the phreatic line at the left side of the flow domain coincides with the reservoir level at all time. Water infiltrates into the dam under stable flow conditions and reaches the toe of the embankment in about 4.6 days. Due to the large permeability contrast of the rock fill in the toe of the dam and the soil in the rest of the dam only a small pressure gradient supports the water flux out of the soil through the rock toe and the saturated zone in the toe is hardly visible. This situation generates an internal seepage face, which is most clearly present in the steady state situation. Due to the flow in a small zone above the phreatic level, the pressure will not be zero at the material interface and the solution does not match the saturated model solution exactly. However, in the physical system the pressure at the material interface will be less than zero as well. The actual values depend on the unsaturated soil behavior. Figure 5.2 compares results found by the adaptive multilevel method with results obtained by a fine mesh

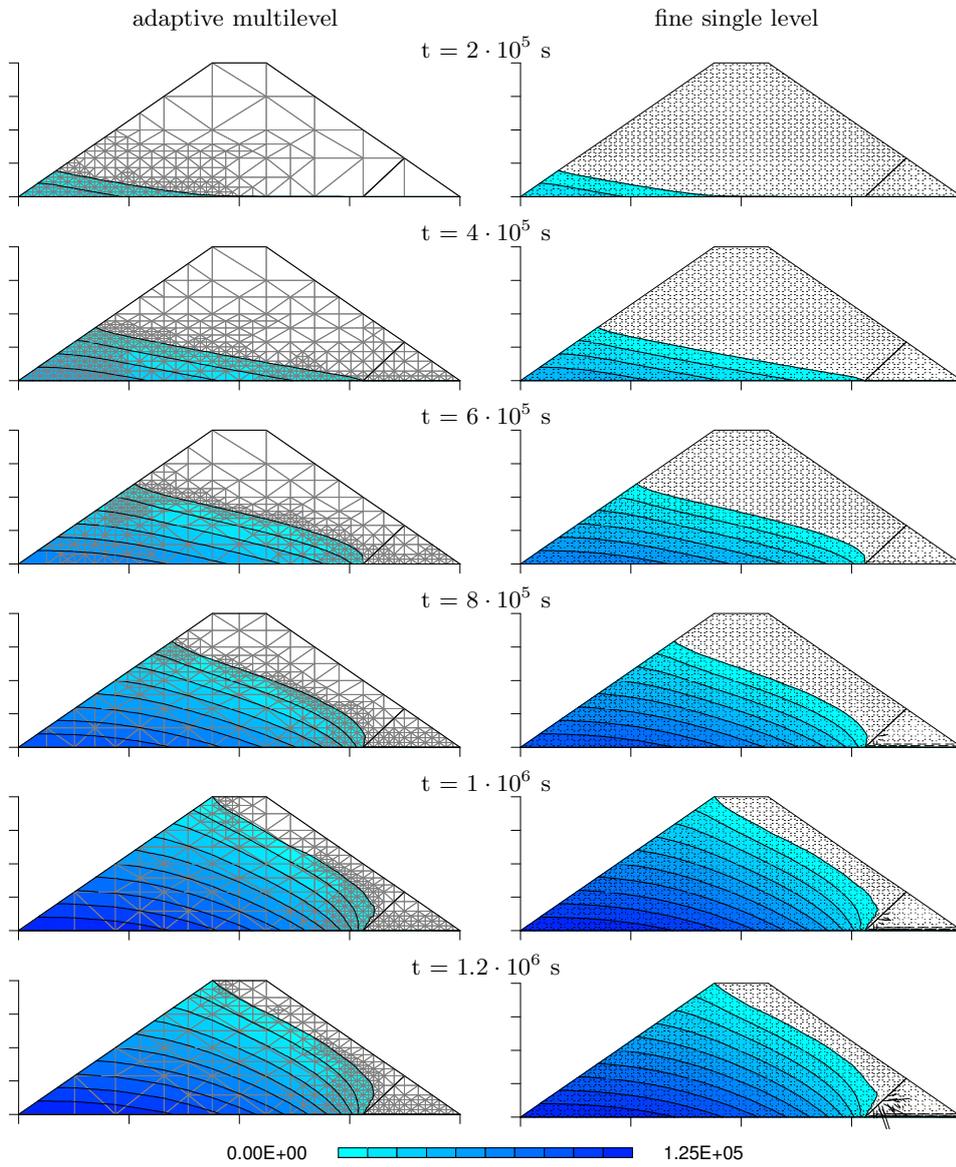


Figure 5.2: Pressure (Pa) fields for Lamb and Whitman's problem.

computation. This mesh applies level four elements over the entire flow domain. The level four elements are shown in parts of the adaptively refined mesh. Results on the adaptive mesh compare well with results on the fine mesh. Mesh refinement corrects the error in the continuity of flux over the element boundaries. For this simulation the refinement concentrates on the phreatic line.

The computational behavior for the Lamb and Whitman problem is presented by figure 5.3. This figure shows the number of nodes and the related pressure field error as a function of time. The error is calculated as

$$\|e\|_2 = \sqrt{\sum_{e=1}^{n_e} \int_{\Omega_e} (N_a \tilde{u}_a - N_a u_a)^2 d\Omega}. \quad (5.6)$$

where  $\tilde{u}_a$  denotes a fine scale reference solution. Scaling the error by  $\|\tilde{u}\|_2$  could reduce the dimension of the parameter.

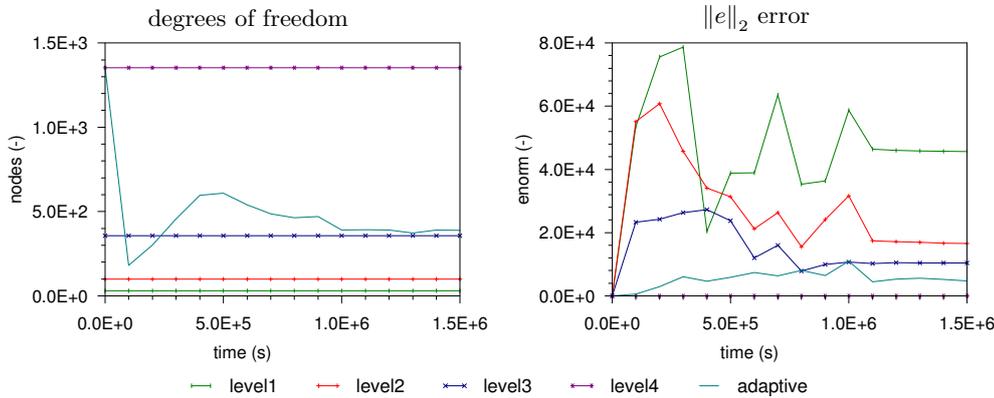


Figure 5.3: Computational results for Lamb and Whitman's problem.

The adaptive mesh solution reduces the number of nodes from 1353 on level 4 to about 357 on level 3. Still the error norm of the adaptive algorithm is well below the norm for a level 3 computation over the entire simulation. The enhanced accuracy of the adaptive algorithm at  $5 \cdot 10^5$  s requires the generation of a limited number of additional elements. These elements are positioned mainly along the phreatic line. As the length of the phreatic line decreases later on, the number of additional elements is reduced as well.

### 5.1.2 Forsyth et al's problem

Forsyth et al. simulated the infiltration process into a large caisson in a two dimensional cross section [50]. The horizontal dimension of the caisson is 8 m and vertical dimension equals 6.5 m. The problem shows a gradual change in permeability over four materials, which are present in the caisson. A zone of  $2 \times 1$  m, with a permeability of 41.52 m/d is included in a larger zone of  $8 \times 5.6$  m with a relatively low permeability of 4.15 m/d. On top of this zone rests a layer of 0.5 m with a permeability of 4.70 m/d. The layer on top of this layer is 0.4 m thick and has a permeability of 7.91 m/d. Figure 5.4 shows the layout of the problem posed by Forsyth et al [50].

The simulation, presented here, follows the benchmark problem outlined in the Fe-flow manual [40]. The manual considered two initial field conditions. These conditions pose a fully unsaturated non-equilibrium starting condition. The infiltration process

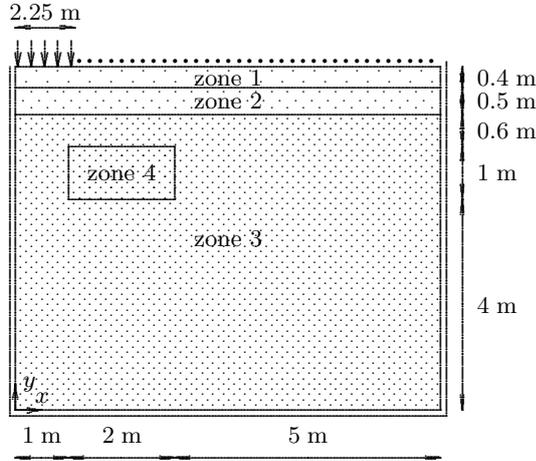


Figure 5.4: Forsyth et al.'s problem definition.

was simulated over 30 days. A time step of  $10^{-3}$  days, and a spatial discretization by  $90 \times 21$  quadrilateral elements was selected. Saturation results were compared for a mesh of 28,917 nodes which was obtained by splitting each quadrilateral element into two triangular elements and then refining the triangular elements twice.

This section selects the most extreme initial case for comparison. The hierarchy of finite element meshes consists of 99, 357, 1353, 5265 nodes, and 162, 644, 2568, 10,256 elements. At the coarsest level, the size of the elements range from 0.5 m up till 1.0 m in x-direction and 0.4 m till 1.0 m in y-direction. The simulation period of  $2.592 \cdot 10^6$  s was divided in 300 steps of 8640 s each.

Initial dry conditions pose a constant pressure on the flow domain, and read

$$p^0 = -10^6 \text{ Pa.} \quad (5.7)$$

These conditions do not state an equilibrium situation, so the result of the time dependent simulation will be influenced by the starting conditions.

An infiltration boundary condition applies to the left most 2.25 m on the top of the model, according to

$$q_n = 0.02 \text{ m/d.} \quad (5.8)$$

Feflow sets the remaining part of the boundaries as impervious. This condition is simply expressed by  $q_x = 0$  on the vertical boundary parts, and  $q_y = 0$  on the horizontal boundary parts. The proposed boundary conditions do not state a well-posed boundary value problem, because the actual value of the pressure solution depends on the initial field only. A boundary value problem is well posed if it satisfies the compatibility condition and it has a unique solution. Well-posed problems are solvable. Compatibility requires that the sum of all source terms is zero. Additional Dirichlet conditions enforce a unique solution. The simulation presented here applies alternative boundary conditions at the top of the domain, which read  $p = -10^6$  Pa.

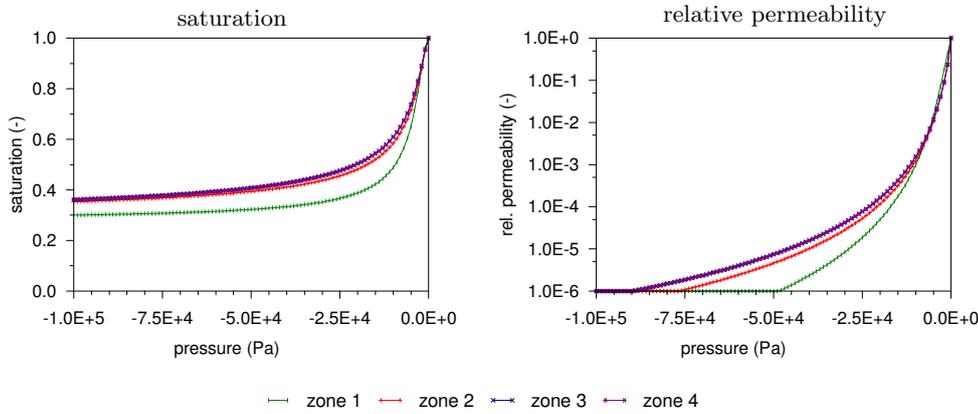


Figure 5.5: Material behavior for Forsyth et al.'s problem.

The density of the water  $\rho$  is  $1000 \text{ kg/m}^3$ , and its viscosity  $\nu$  is  $10^{-3} \text{ Pa s}$ . The gravity constant is set to  $10.0 \text{ m/s}^2$ . Figure 5.5 shows Van Genuchten material behavior for Forsyth et al.'s problem graphically. Table 5.2 includes the coefficients for these Van Genuchten model functions.

Table 5.2: Van Genuchten shape factors for Forsyth et al.'s problem.

zone	$S_r$	$n$	$K$ m <sup>2</sup>	$g_a$ 1/m	$g_l$	$g_n$
1	0.2771	0.368	$9.153 \cdot 10^{-12}$	3.34	0.5	1.982
2	0.2806	0.351	$5.445 \cdot 10^{-12}$	3.63	0.5	1.632
3	0.2643	0.325	$4.805 \cdot 10^{-12}$	3.45	0.5	1.573
4	0.2643	0.325	$4.805 \cdot 10^{-11}$	3.45	0.5	1.573

Figure 5.6 shows the saturation fields for a number of time steps, and compares the results of the adaptive computation to the fine mesh calculation. The saturation fields are derived from continuous pressure fields. The saturation fields show discontinuities over the material boundaries due to the difference in unsaturated behavior for the individual soil materials. Adaptive grid results compare well with fine scale results, although the adaptive algorithm reduces the number of elements considerably. Grid refinement concentrates on zones with a sharp saturation gradients and zones with curved streamlines. Both zones were detected by considering the error in continuity of flux over the element boundaries. The adaptive simulation starts with a very coarse mesh. After one day of infiltration the infiltration front reaches the second layer. During this first period infiltration is almost vertical. Next water infiltrates into the first layer in horizontal direction as well because of the permeability contrast with the second layer. The infiltration speed is lower in the second layer mainly due to the

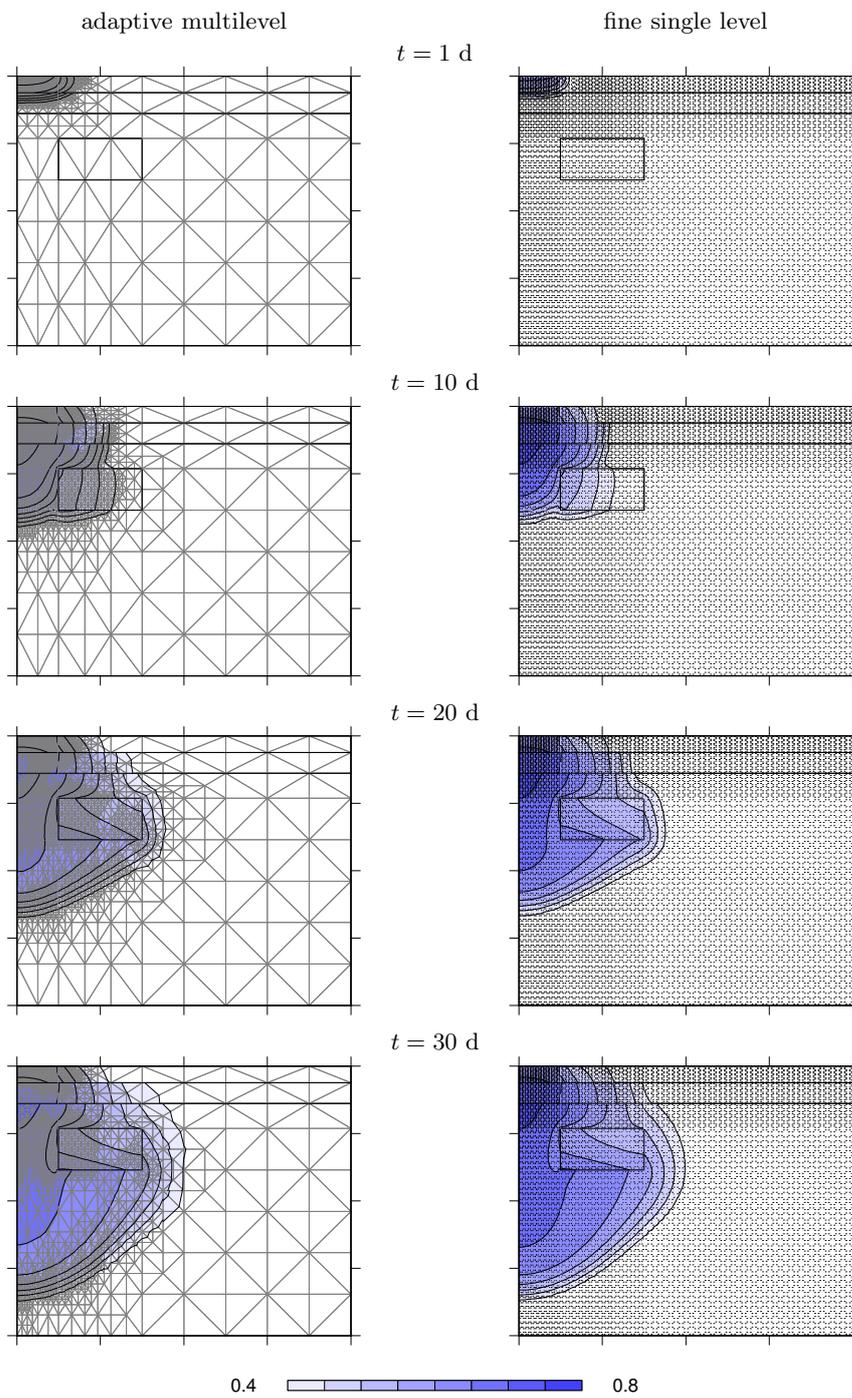


Figure 5.6: Saturation (-) fields for Forsyth's problem.

lower value of intrinsic permeability of this layer. The degree of saturation in the top layer is less as a result of the difference in saturation curves, which apply to both layers. After 10 days of infiltration the infiltration front extends to the bottom level of the included zone. Saturation contours are continuous over zone 3 and 4, because the unsaturated behavior of the materials that fill these zones is the same. The intrinsic permeability of zone 4 is a factor 10 higher than the intrinsic permeability of zone 3. During the infiltration process the embedded zone drains its surrounding and transports the incoming water faster than the material next to it. As a result bypass flow occurs that infiltrates zone 3 at the bottom of zone 4 after some time.

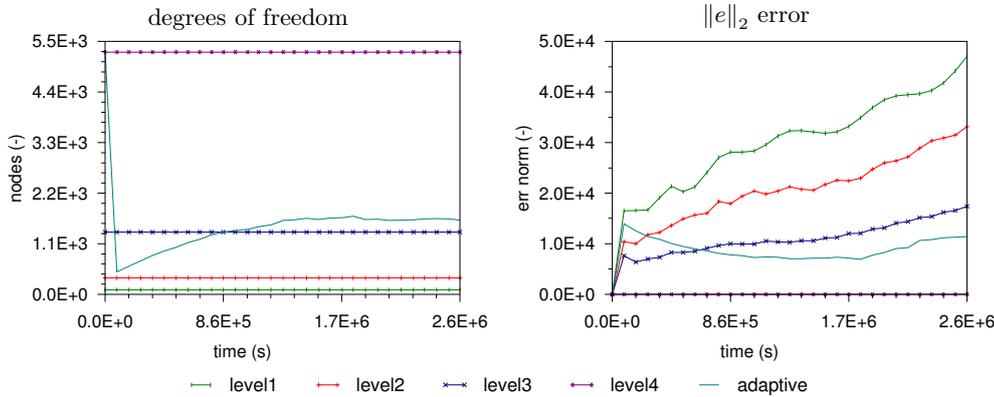


Figure 5.7: Computational results for Forsyth et al.'s problem.

Computational results for the modified Forsyth et al.'s problem are shown in figure 5.7. The accuracy of the pressure field was calculated by equation (5.6). In the first period of the simulation, up to one day, the number of nodes in the adaptive mesh simulation is well below level 3 refinement of the mesh. This gives less accurate results. From that moment on the number of nodes in the adaptive mesh corresponds approximately to the level 3 mesh, which was constructed out of 1352 nodes and 2568 elements. The infiltration front however, was captured by level 4 elements, improving the accuracy of the adaptive mesh solution. A full level 4 mesh would require 5265 nodes and 10256 elements.

## 5.2 Solute transport and heat transport

This section deals with solute transport and heat transport and covers the verification of the adaptive formulation for two well-known fully coupled flow and transport benchmarks: Henry's problem and Elder's problem. These cases involve saturated groundwater flow through homogeneous domains. Henry adds mass transport to the problems, and Elder couples heat transport to the flow problem. Both problems consider density changes and density driven flow. The problems are closed by prescribed fluid flux conditions or prescribed pressure conditions. The boundary conditions either generate flow into the domain or flow out of the domain. For the pressure condition this

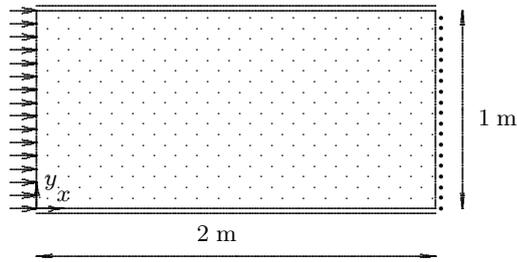


Figure 5.8: Definition of Henry's problem

depends on the pressure in the domain. Flow into the domain constitutes a prescribed solute mass flux or a prescribed energy flux. A weak boundary condition simulates the solute transport or heat transport out of the domain, by imposing a zero diffusive flux.

Section 5.2.1 presents Henry's problem. This problem involves salt water intrusion into a domain in the opposite direction of the natural fresh water flow [4, 21, 39, 53, 66, 83, 79, 90, 107]. Flow conditions for this problem are stable as salt water under lays fresh water. The problem verifies the accuracy and efficiency of mesh refinement for this type of density driven flow, and focuses on geometrically nonlinear boundary conditions that result from the prescribed pressure conditions, which apply on the outflow side of the domain. Section 5.2.2 presents Elder's problem [47, 4, 83, 107, 121, 39, 79, 66]. This problem involves potentially unstable heat flow and considers a domain which is heated from below. This benchmark focuses on flow that is purely driven by density differences and generates natural convection. Special attention will be paid to the effect of the discretization of the problem on instabilities.

### 5.2.1 Henry's problem

Henry's problem involves seawater intrusion into a confined layer. At the inland side fresh water flows into the domain, and salt water intrudes the domain at the opposite side. Due to the difference in density, fresh water flows over the salt water. Henry derived an analytical expression for the stream function and calculated the concentration field in the form of a Fourier series. The resulting algebraic equations, which determine the coefficients of the Fourier series were solved numerically. Henry obtained a dimensionless solution for the steady state solution in the cross section.

The simulation presented here, follows the benchmark problem outlined in the Sutra manual [118]. In this manual Voss proposed dimensions for the problem that make a comparison with other models possible. He also extended the original problem to a transient problem, which finally obtains the steady state situation. The total simulation time is 100 minutes, for which the steady state can essentially be reached. The cross section has a length of 2.0 m and a height of 1.0 m. The Sutra manual proposed a mesh that consists of  $20 \times 10$  elements of size  $0.1 \times 0.1$  m. A uniform time step of 1.0 minute was selected for the discretization in time. Figure 5.8 shows the layout of Henry's problem.

The Feflow manual proposes a fine scale elements size of  $0.0125 \times 0.0125$  m. Based on this information, the hierarchy of finite element meshes constructed here, consists of 66, 231, 861, 3321 nodes, and 105, 410, 1620, 6440 elements. At the coarsest level the size of the elements equals 0.2 m in both directions. The simulation period of  $6.0 \cdot 10^3$  s was divided in 100 steps of 60 s each.

The initial state of the system is obtained by a steady state computation based on freshwater inflow, zero concentration on both vertical boundaries, and specified pressures at the sea side. This pressure condition already corresponds to enforcing a hydrostatic seawater pressure distribution as will be shown below.

No-flow boundary conditions apply to the top and bottom of the domain. Fixed-flow conditions apply to the left vertical boundary. Here a mass flux of  $6.6 \cdot 10^{-2}$  kg/s holds and the salt mass fraction is set to zero. This condition poses the inflow of fresh water into the domain, and is written as

$$q_x = 6.6 \cdot 10^{-5} \text{ m/s} \quad \cap \quad \bar{\omega} = 0. \quad (5.9)$$

At the right vertical boundary, a hydrostatic pressure is prescribed. Water that enters the domain through the lower part of the boundary, has the density of seawater. The diffusive influx  $j_n = -D\partial\omega/\partial x$  of the water that leaves the domain at the upper part of the boundary is set to zero. The extend of both boundary parts varies throughout the simulation. This geometrically nonlinear boundary condition reads

$$p = \rho_s g(h - y) \quad \cap \quad j_n = \begin{cases} 0 & \text{if } q_x > 0 \\ (\bar{\omega} - \omega) q_x & \text{if } q_x \leq 0 \end{cases}, \quad (5.10)$$

where  $\bar{\omega} = 3.571 \cdot 10^{-2}$ .

Table 5.3 lists the material parameters and their values. The density of the binary fluid depends on the mass fraction of the solute. The constitutive relation for the density is simplified to

$$\rho = \rho_0 + \beta_\omega (\omega - \omega_0). \quad (5.11)$$

Henry's problem specifies a velocity-independent diffusion. This Fickian dispersion is incorporated into the model by setting the molecular diffusion to  $18.8571 \cdot 10^{-6}$  m<sup>2</sup>/s. The total dispersion coefficient then equals  $6.6 \cdot 10^{-6}$  m/s<sup>2</sup>, as the porosity is set to 0.35. In physical systems a velocity based dispersion describes the mixing process more accurately, and the transition zone between salt and fresh water can be relatively thin. The gravity constant is 9.8 m/s<sup>2</sup>.

Figure 5.10 presents mass fraction fields for a number of time steps. Initially the domain is filled with fresh water only. During the simulation salt water intrudes into the domain. The greater density of the salt water drives the intrusion. At the bottom of the flow domain the saltwater flux points in the opposite direction of the fresh water flux at the top of the domain. As a result a single eddy will evolve. At the top of the eddy mixing occurs. Both adaptive mesh results and fine grid results compare well, and the final result at 1.6 hours corresponds to the steady state result presented in the Sutra manual.

Figure 5.9 shows the number of nodes of the adaptively refined mesh and compares them to the number of nodes on the uniform meshes. The corresponding accuracy of

Table 5.3: Material parameters for Henry's problem.

parameter	value
permeability skeleton $\kappa$	$1.020408 \cdot 10^{-9} \text{ m}^2$
density fresh-water $\rho_0$	$1000 \text{ kg/m}^3$
density sea-water $\rho_s$	$1024.99 \text{ kg/m}^3$
density variability coefficient $\beta_\omega$	$700 \text{ kg/m}^3$
viscosity water $\nu$	$10^{-3} \text{ Pas}$
porosity skeleton $n$	0.35
diffusion coefficient $D_c$	$1.88571 \cdot 10^{-5} \text{ m}^2/\text{s}$
longitudinal dispersivities $\alpha_l, \alpha_t$	0 m
compressibility skeleton $\alpha_p$	$0 \text{ m}^2/\text{N}$

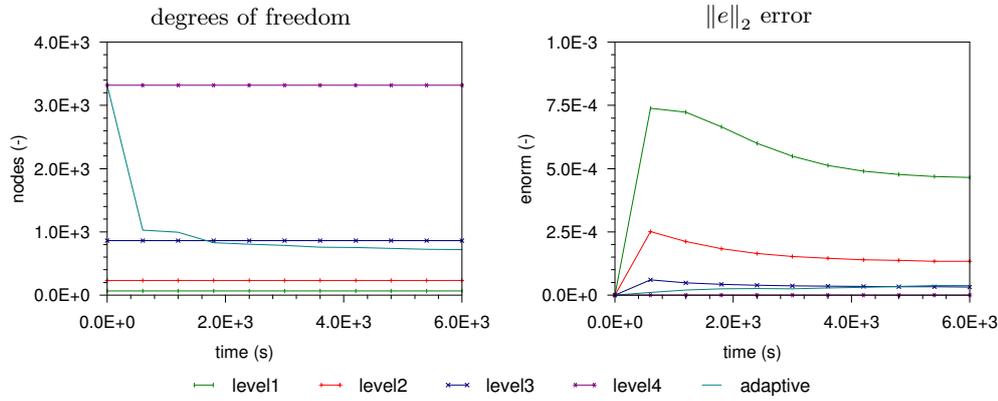


Figure 5.9: Computational results for Henry's problem.

the solution, now expressed in mass fraction according to equation (5.6), is also given as a function of time. The figure shows that the number of nodes in the adaptive mesh computation approximately corresponds to the number of nodes in the level 3 computation. The accuracy of both level 3 calculation and adaptive mesh computation only slightly differs from the level 4 computation. Aiming at level 3 accuracy with an adaptive mesh build out a number of nodes that corresponds to the level 2 mesh could improve the efficiency of the adaptive method. Figure 5.10 shows a large part of the domain that contains level 1 elements. This points out that adaptive refinement of local phenomena is especially efficient.

### 5.2.2 Elder's heat problem

Elder's heat problem considers transient thermal convection in a confined cross-sectional rectangular region. Part of the bottom boundary holds a fixed temperature greater

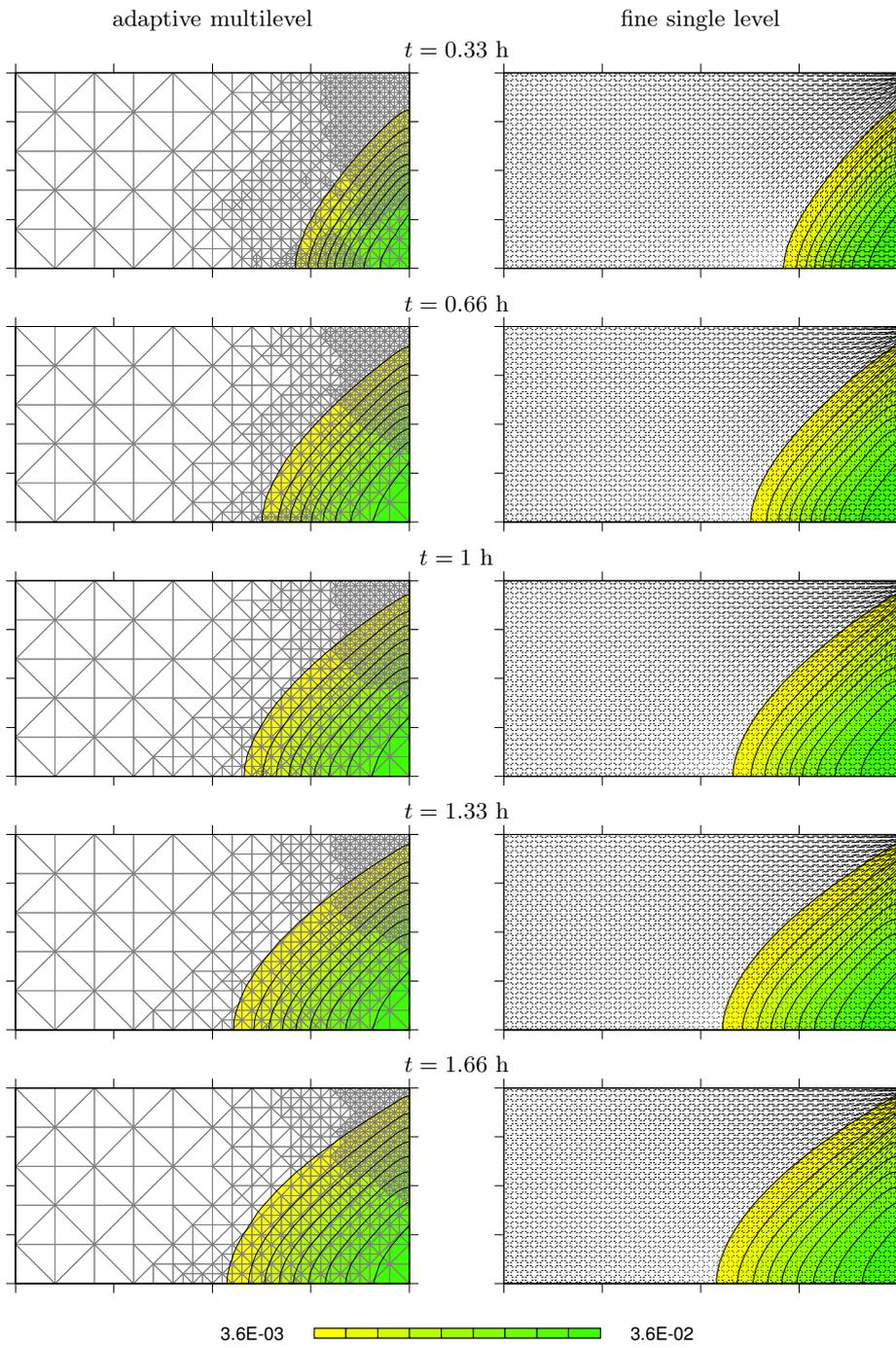


Figure 5.10: Mass fraction (-) fields for Henry's problem.

than the initial uniform temperature in the domain. The conducted heat into the pore water decreases the density and induces circulating flow. Elder used a finite difference technique to solve the set of equations, which was formulated in terms of dimensionless variables.

The simulation that is presented here, follows the benchmark problem outlined in the Hst3d manual [79]. Figure 5.11 shows the layout of the modified Elder's problem. The cross section is 4 m long and 1 m high. The heat source is present along the middle 2 m of the bottom boundary. The box is closed by impermeable walls on the remaining horizontal and vertical boundaries. Kippe [79] proposed a uniformly discretization of

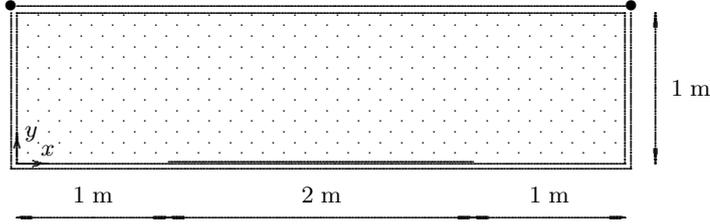


Figure 5.11: Definition of Elder's thermal problem.

the domain with nodal spacing of 0.091 m in the horizontal direction and 0.040 m in the vertical direction, and uniform time steps of 596 s over a period of  $1.1925 \cdot 10^5$  s. This period corresponds to a dimensionless time of 0.1, used by Elder.

The hierarchy of finite element meshes consists of 66, 231, 861, 3321 nodes, and 105, 410, 1620, 6440 elements. At the coarsest level, the size of the elements equals 0.2 m in x-direction and 0.2 m in y-direction. The simulation period of  $1.192 \cdot 10^5$  s was divided by 2000 steps of 59.6 s each.

Initial conditions pose a hydrostatic pressure and a uniform temperature on the flow domain. This initial condition reads

$$p^0 = \rho_0 g(h - y) \quad \cap \quad T^0 = 283 \text{ K.} \quad (5.12)$$

The boundary condition at the top of the domain reads

$$q_y = 0 \quad \cap \quad T = 283 \text{ K.} \quad (5.13)$$

A pressure of 0 Pa at the upper corners of the region is proposed in order to generate a well posed problem. At the corner points of this boundary, the condition is replaced by

$$p = 0 \quad \cap \quad T = 283 \text{ K.} \quad (5.14)$$

In the middle part of the boundary at the bottom, the following conditions hold

$$q_y = 0 \quad \cap \quad T = 293 \text{ K.} \quad (5.15)$$

Table 5.4: Material parameters for Elder's problem.

parameter	value
thermal conductivity solid phase $\lambda^s$	2.0 W/mK
thermal conductivity fluid phase $\lambda$	0.6 W/mK
thermal heat capacity solid phase $c\rho^s$	$2.0 \cdot 10^6$ J/m <sup>3</sup> K
specific heat capacity fluid phase $c$	4182 J/kgK
thermal expansion coefficient fluid phase $\beta_T$	$2 \cdot 10^{-4}$ 1/K
reference density fluid phase $\rho_0$	1000 kg/m <sup>3</sup>
reference temperature $T_0$	293 K
density solid phase $\rho^s$	2650 kg/m <sup>3</sup>
permeability skeleton $\kappa$	$9.084 \cdot 10^{-9}$ m <sup>2</sup>
porosity skeleton $n$	0.1
viscosity fluid phase $\nu$	$10^{-3}$ Pas
molecular diffusion coefficient $D_c$	0 m <sup>2</sup> /s
dispersivities porous medium $\alpha_l, \alpha_t$	0 m
compressibility skeleton $\alpha_p$	0 m <sup>2</sup> /N

At the remaining part of the bottom boundary, and on the vertical boundaries the conditions read

$$q_y = 0 \quad \cap \quad \frac{\partial T}{\partial y} = 0, \quad q_x = 0 \quad \cap \quad \frac{\partial T}{\partial x} = 0. \quad (5.16)$$

Table 5.4 presents the material parameters for Elder's problem and their values. The fluid density varies between 1000 kg/m<sup>3</sup> and 998 kg/m<sup>3</sup>. The constitutive relation for the density is simplified by

$$\rho = \rho_0 - \beta_T (T - T_0). \quad (5.17)$$

The compressibility of water and solids is disregarded. The gravitational constant equals 9.807 m/s<sup>2</sup>.

Figure 5.12 presents the computed temperature fields for an adaptively refined mesh and the corresponding fields for a uniform fine mesh. Both series only show half of the original flow domain, and the results will be discussed for this reduced domain. In the first stage of the simulation two eddies emerge on both sides above the heater end, which are driven by the rising fluid. On the fine grid two small eddies become visible at 3.3 h. These eddies are positioned next to the larger eddies. A third small eddy becomes visible at 6.6 h at the right vertical side of the domain, which models symmetry of the original flow domain. The adaptive mesh shows different results. Here convective transport of heat triggers an eddy at the right side of the domain at an early stage. At 6.6 h the system contains five eddies of about equal strength. For both fine scale computation and adaptive mesh computation the eddies shift to the right later

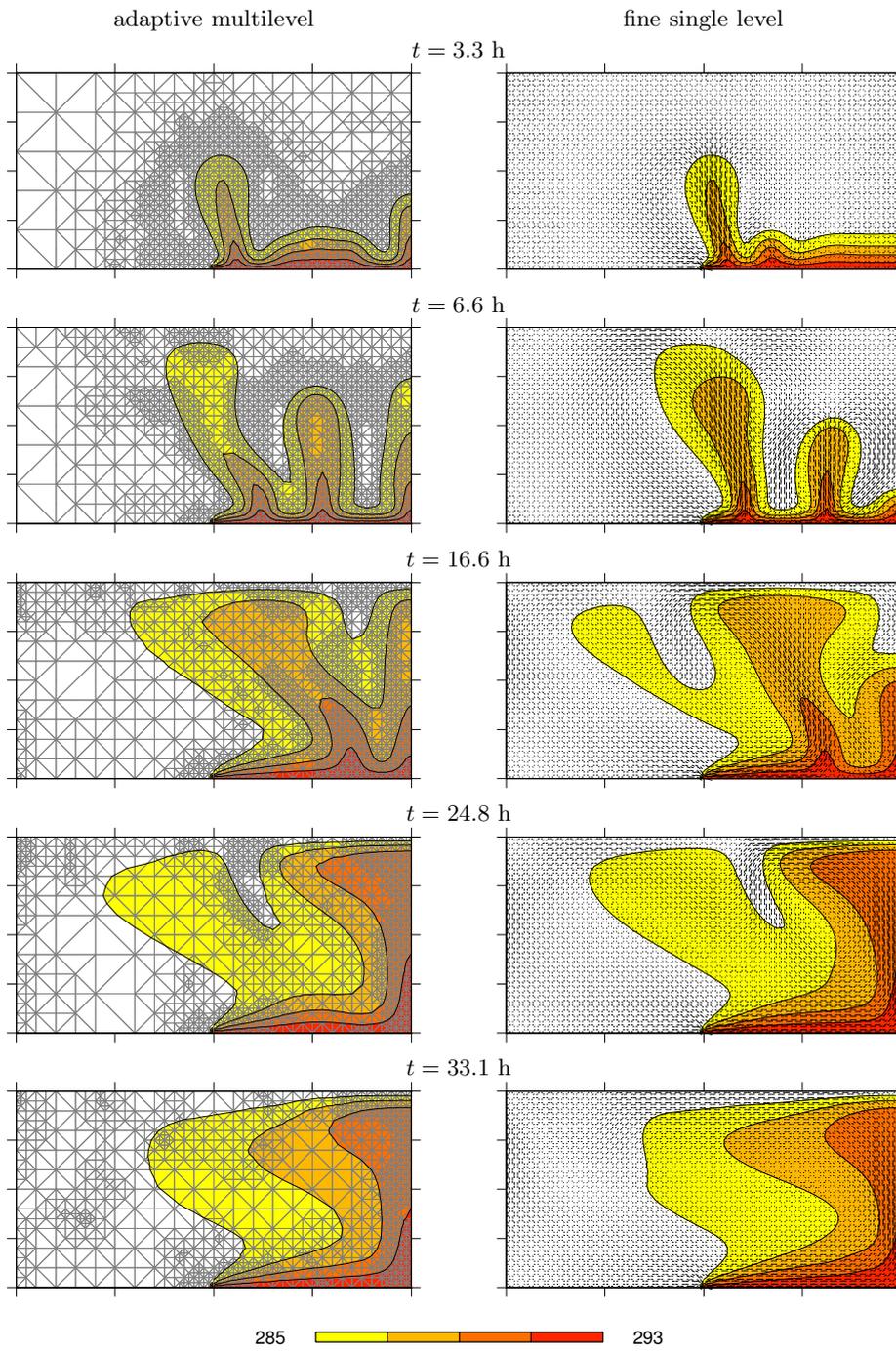


Figure 5.12: Temperature (K) fields for Elder's problem.

on and form a system that contains three eddies at 16.6 h. Finally the eddies melt into one roll and reach a stable steady state situation. In the center part of the original domain the fluid flow is upward, at the ends downward flow occurs. The figure shows a large mesh dependency of the results in the development stage.

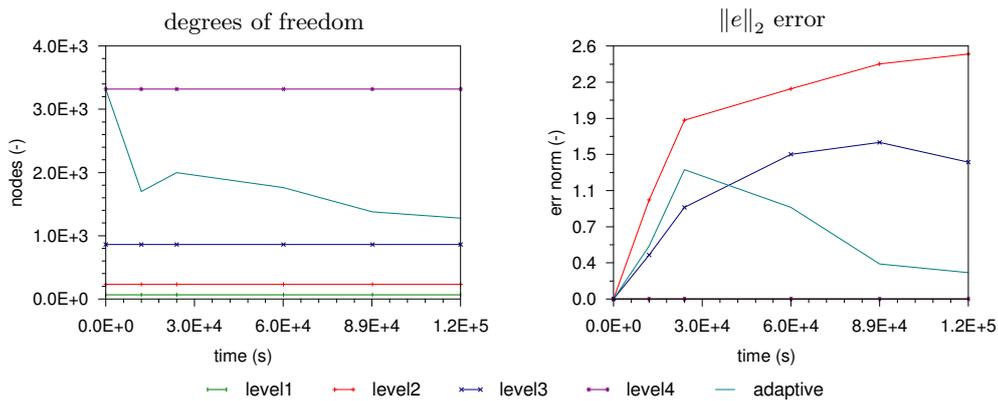


Figure 5.13: Computational results for Elder's problem.

Figure 5.13 presents the adaptively generated number of nodes and the accuracy of the temperature field as a function of time. Although the number of nodes is above the number of nodes generated by a level 3 discretization, the accuracy of the adaptive mesh computation does not improve much during the early development stage. The reason is that the adaptively refined mesh triggers a different convective flow pattern. At the end of the computation the adaptive mesh computation gives more accurate results than the level 3 calculation although the number of nodes is about the same.

Elder's problem poses a complicated verification test, as the size of the mesh influences the simulated transport mechanism. Adding a local refinement criterion on the projected error of the mass fraction might resolve this problem. At this stage the flow problem has been solved on a hierarchy of adaptively refined finite element meshes using the proposed multigrid solver. The transport equation has been solved by a preconditioned bi-conjugate gradient solver on a single level. Solving the transport equation sequentially over at least two levels facilitates extrapolation to the finest level where the refinement criterion could be postulated.

# Chapter 6

## Flow simulation

The tenth SPE (Society of Petroleum Engineers) comparative solution project aimed to compare upgridding and upscaling techniques for two problems [26, 56, 86]. The first problem involved flow through a two-dimensional vertical cross section, for which the geological model was captured by two-thousand cells. The model domain is  $2500 \times 50$  ft, and a fine grid covered the model by uniform sized grid blocks of  $100 \times 20$  ft. Upscaling techniques should generate equivalent permeability for a specified uniform upgridded model of 25 coarse grid cells and for a non-uniform adaptive grid model of 100 cells at maximum. The second problem involved a three-dimension waterflood problem of over 1.1-million cells. For this number of cells results could still be compared to fully resolved computations. The model domain is  $1200 \times 2200 \times 170$  ft. The top 70 ft represents the near-shore Tarbert formation, and the bottom 100 ft represents the fluvial Upper Ness formation. For this model the fine-scale cell size is  $20 \times 10 \times 2$  ft. Porosity and permeability are constant over the fine scale cells. The porosity field is strongly correlated to the permeability field. The quotient of vertical and horizontal permeability is 0.3 in highly permeable channels and 0.001 in the background formation. Both formations show large permeability variations of 8 till 12 orders of magnitude. Participants to the project produced upscaled permeabilities on grids of  $30 \times 55$  cells up till  $5 \times 5$  grid cells. The three-dimensional reservoir is produced using a water drive from a single well in the center of the model and four producers in the corners of the model. Wells are completely vertical throughout the model. Prescribed bottom hole pressure hold for all wells. The process was simulated over 2000 days of production. Production curves and average field pressure were compared.

This chapter includes two applications of the adaptive multiscale finite element method on these highly heterogeneous domains. Instead of oil production, energy storage and heat production will be considered here. Section 6.1 considers geothermal storage in the shallow marine formation. Section 6.2 focuses on geothermal energy production in a deep fluvial system. For both applications the geological model and the corresponding flow model will be presented. Adaptive multiscale results will be compared with fully resolved solutions.

## 6.1 Marine system

The top 70 ft of the second SPE model represents the near-shore Tarbert formation. The distribution of porosity and permeability was given for  $60 \times 220 \times 35$  cells. The parameter values are constant for each cell and discontinuous over the cell boundaries. Here the problem is reduced to  $60 \times 220 \times 32$  cells, because this fits a hierarchy of three levels. The geological model now includes 422,400 cells. The size of the problem was set to  $366 \times 671 \times 20$  m, and the size of the cells equals  $6.1 \times 3.05 \times 0.625$  m. The aspect ratio of the elements is 9.76, which is relatively large. The flow model splits each quadrilateral cell into 6 tetrahedral elements and generates 2,534,400 elements in total. The number of nodes equals 444,873, and as flow and transport will be considered, two systems with 444,873 degrees of freedom apply. The coarsened mesh at level 2 includes 52,800 cells, and the level-one mesh contains 6,600 cells. For this mesh, the degrees of freedom reduce to 16,128.

This section presents flow simulation results for a two-dimensional area: the top layer of the formation. The geological model is restricted to  $60 \times 220$  cells on the finest level. The number of elements in the hierarchy of finite element meshes is 13,200, 3,300 and 825, the number of nodes equals 13,481, 3,441, and 896.

### 6.1.1 Geological model

Figure 6.1 shows the porosity field for the geological model of the Tarbert formation, and Figure 6.2 presents the distribution of the porosity and the permeability for the

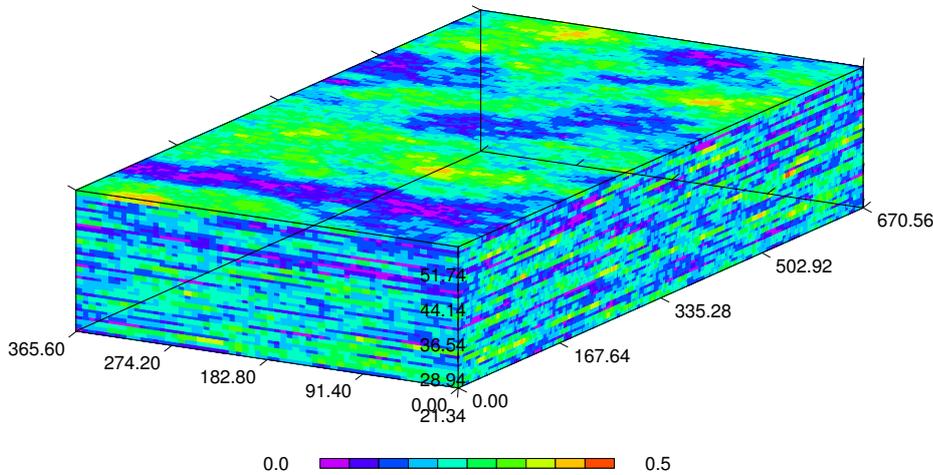


Figure 6.1: Tarbert porosity (-) field.

top layer of the marine system. Porosity varies between 0.05 and 0.45, permeability fluctuates between  $3.03 \cdot 10^{-17} \text{ m}^2$  and  $4.65 \cdot 10^{-11} \text{ m}^2$ . Both parameters fit into a single distribution.

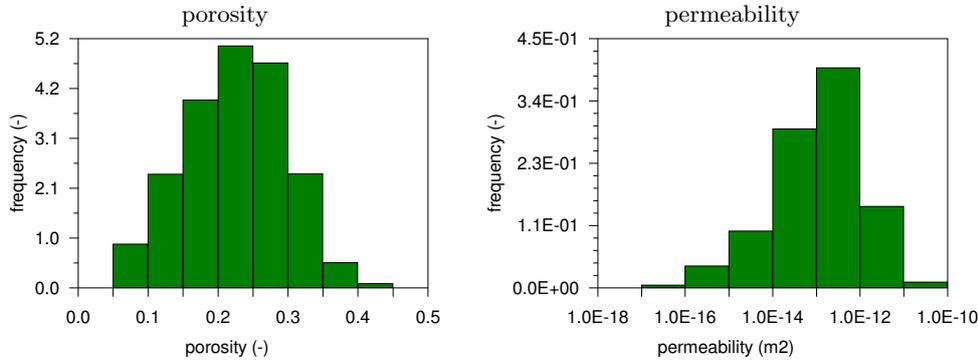


Figure 6.2: Material property distribution Tarbert formation top layer.

Figure 6.3 presents the porosity distribution and the permeability distribution over the top layer. The figures show the small scale distribution on level 3, and its upscaled values at level 2 and level 1. The pictures show that as the scale decreases, the resolution becomes less. The pictures at the bottom presents adaptive scaling results. The refinement is based on the flow problem imposed. Here, the error in volumetric flux over the element boundaries, as discussed in chapter 4, detects the regions on which the mesh has to be refined. The pictures show a low permeable zone at the left side that stretches from the bottom to the top of the flow domain. This zone is present at all scales.

### 6.1.2 Flow model

The flow model simulates a simplified heat storage system. In the summer period, a source injects heated water into the system. During this period a sink produces water at subsurface temperature for cooling purposes. In winter time the system is reversed. During this period the energy stored in the system is recaptured and cooled down water is injected.

The simulation considers the top layer of  $607 \times 366$  meter of the marine system. The fine scale permeability distribution was first multiplied by a factor of ten in order to represent an aquifer, then the porosity was set to 0.35. Natural conditions for this layer impose a pressure of  $9 \cdot 10^5$  Pa, as the layer is positioned 90 meter below the phreatic surface, and a temperature of  $13^\circ$  C. Artificial conditions apply on the boundaries of the flow model. In a top view, the conditions at the vertical boundaries prescribe a constant pressure of  $9 \cdot 10^5$  Pa. Along the horizontal sides of the flow domain no-flow conditions apply. The simulation introduces a hot well at the point 240 - 240 m and a cold well at 480 - 96 m. The injector well and producer well operate independently as their distance is 280 m. The hot well injects water of  $24^\circ$  C in the summer period. The pressure at the bottom of the well is  $1.1 \cdot 10^6$  Pa (20 m above the natural head) and the infiltration rate equals  $1.7 \text{ m}^3/\text{h}$  per meter filter length. Hot water is produced in the winter at the same rate by lowering the head 20 m below the natural head, imposing a pressure of  $7 \cdot 10^5$  Pa. In the winter period the cold well injects water of  $8^\circ$  C into

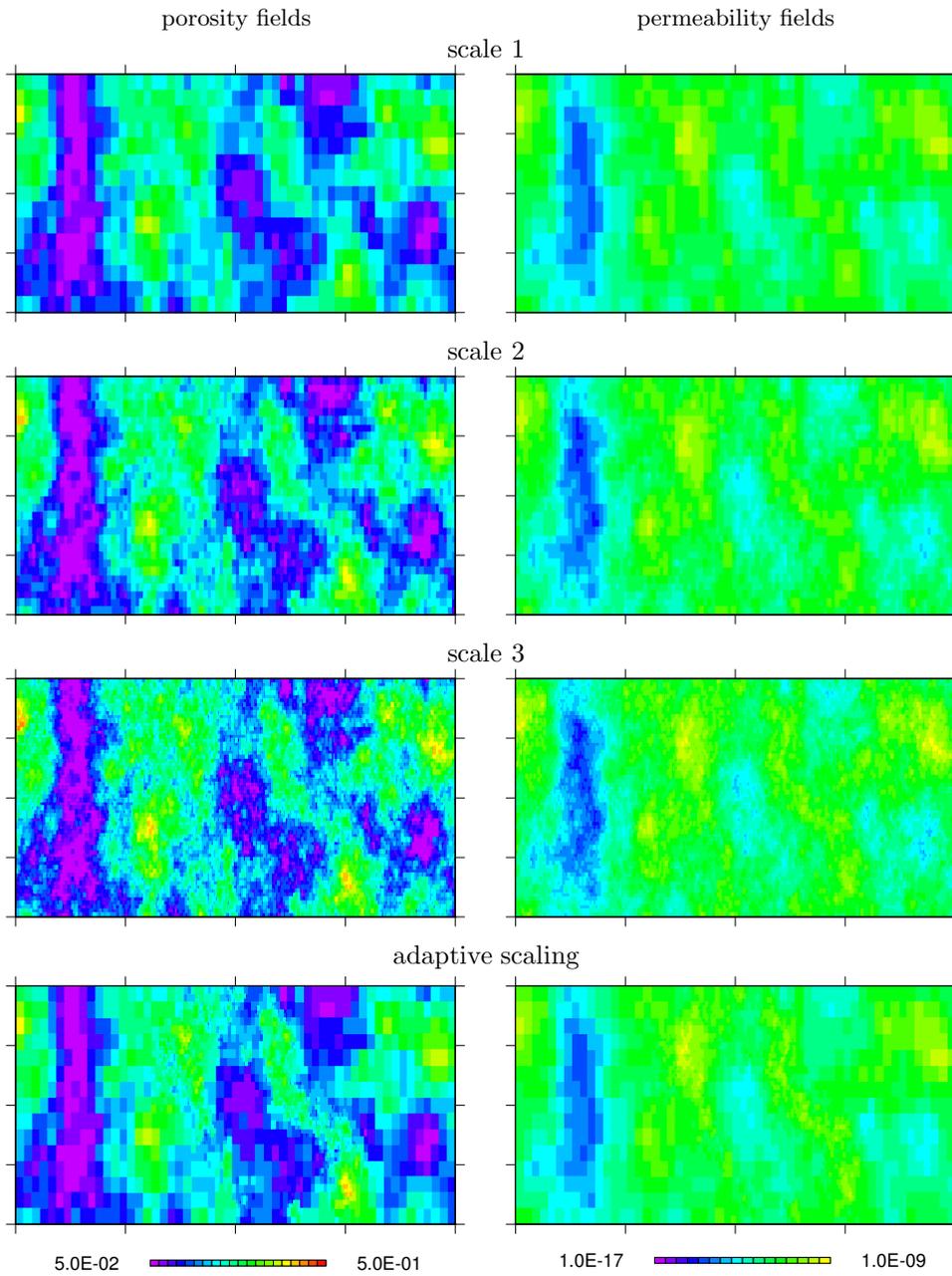


Figure 6.3: Porosity (-) and permeability ( $m^2$ ) fields, heat storage system.

the aquifer at a rate of  $1.7 \text{ m}^3/\text{h}$  per meter filter length, at a pressure of  $1 \cdot 10^6 \text{ Pa}$ . During the summer period the cold well produces water at the same rate by lowering the natural head by 10 m, which sets the producer pressure to  $8 \cdot 10^5 \text{ Pa}$ . A total filter length of 30 m obtains realistic flow rates of  $51 \text{ m}^3/\text{h}$ . The total simulation time is 365 days, divided in 4 seasons of 91.25 days each, and discretized by 400 time steps. The (level 3) finite element mesh was constructed out of 13,481 nodes and 26,400 elements.

Table 6.1 presents the material parameters for the flow simulation and their values. The fluid density is practically constant, as the temperature varies between 281 K and 297 K. The constitutive relation for the density is simplified by

$$\rho = \rho_0 - \beta_T (T - T_0). \quad (6.1)$$

Density effects also play a minor role as a horizontal section has been studied here. The compressibility of water and solid matrix is neglected.

Table 6.1: Material parameters for the flow simulation.

parameter	value
thermal conductivity solid phase $\lambda^s$	1.7 W/mK
thermal conductivity fluid phase $\lambda$	0.6 W/mK
thermal heat capacity solid phase $c\rho^s$	$2.324 \cdot 10^6 \text{ J/m}^3\text{K}$
specific heat capacity fluid phase $c$	4182 J/kgK
thermal expansion coefficient fluid phase $\beta_T$	$2 \cdot 10^{-4} \text{ 1/K}$
reference density fluid phase $\rho_0$	$1200 \text{ kg/m}^3$
reference temperature $T_0$	293 K
density solid phase $\rho^s$	$2800 \text{ kg/m}^3$
viscosity fluid phase $\nu$	$10^{-3} \text{ Pas}$
molecular diffusion coefficient $D_c$	$0 \text{ m}^2/\text{s}$
dispersivities porous medium $\alpha_l, \alpha_t$	3 m
compressibility skeleton $\alpha_p$	$0 \text{ m}^2/\text{N}$

Figure 6.4 presents the pressure fields and the velocity fields for the top layer. The coarse scale pressure field already compares well with the pressure field at the fine scale. The velocity field at the coarse scale is not captured that well at the coarse scale. This will have a large impact on the results of transport simulations. The adaptive refinement algorithm is able to correct the velocity field locally. The algorithm restores the accuracy near the injector well and producer well, where flow velocities are relatively large and flow lines are curved.

The heat storage problem was represented by a fully coupled set of algebraic equations presented in chapter 3 although non-linear effects are small. The flow criterion constructs an adapted mesh, which consists of 2,902 nodes and 5,202 elements. The size of the sets of equations have been reduced by a factor of  $4.64^2$ , the number of coefficients in the sparse matrix reduce linearly. The pressure field was resolved by

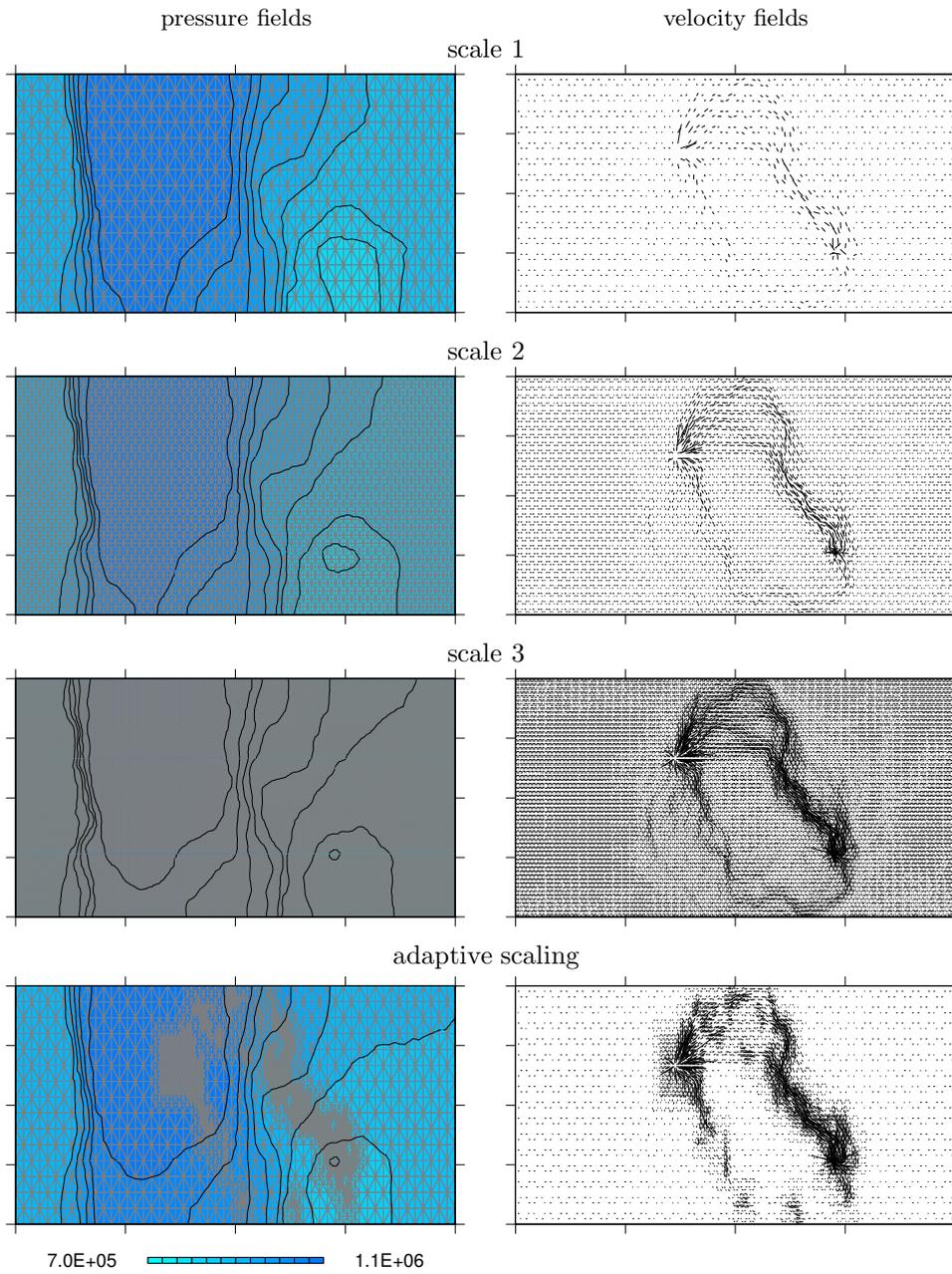


Figure 6.4: Pressure (Pa) and velocity fields, heat storage system.

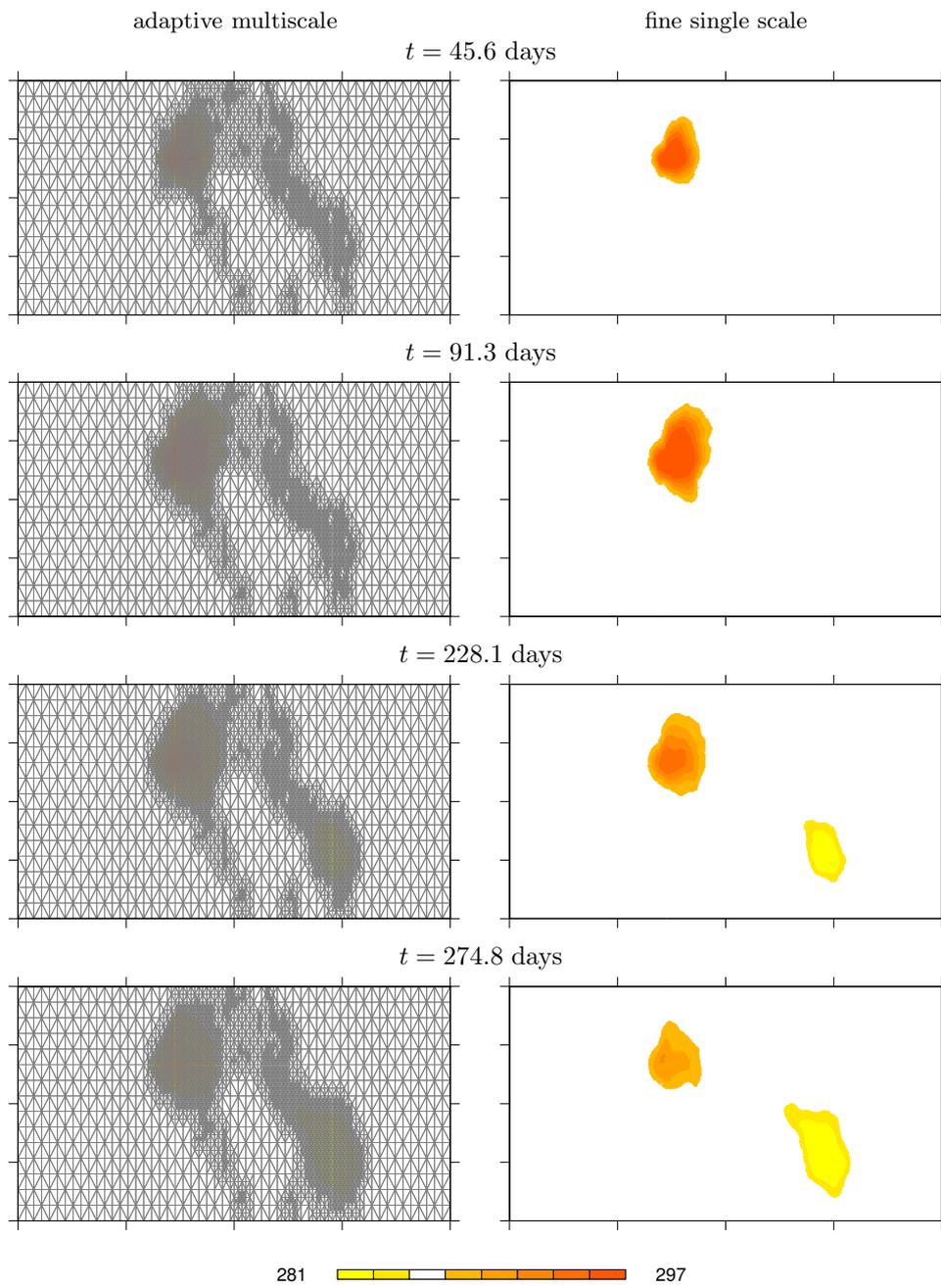


Figure 6.5: Temperature (K) fields, heat storage system.

the multigrid solver outlined in chapter 4. The first step requires 13 cycles with 2 Gauss Seidel pre-smoothing operations and 2 post-smoothing operations. The convergence factor equals 0.22. A single V-cycle resolves the pressure field in the next steps because boundary conditions do not change, storage is neglected, density effects are small, and the mesh does not change. The temperature field was resolved by a pre-conditioned bi-conjugate gradient solver. The computational work for setting up the systems of equations has been reduced by a factor of 5.07, because the velocity field did not change during the simulation and there was no need in constructing a new data structure for the system matrix and the right hand side vector. The overall computation time for the fully refined (level 3) simulation has been reduced by adaptive refinement by a factor 4.72. Posing an additional refinement criterion on local temperature change modifies the mesh per time step and requires the construction of a new system matrix. This reduces the overall computation time factor to 4.17.

Figure 6.5 shows a sequence of temperature fields, and compares results on the adaptively refined mesh to results on a (level 3) fine mesh. The upper pictures at 45.6 days and 91.3 days show a temperature increase during the summer period. The pictures at 228.1 days and 274.8 days show a decrease in temperature during the winter period. The results show that the convective heat flux is relatively low compared to the diffusive heat flux by conduction and hydrodynamic dispersion. The temperature distributions obtained by the adaptive mesh computation compare well with the results of the fine mesh calculation.

Simulations on the Tarbert formation showed that the multiscale finite element method based on an oversampled computation of basis functions sometimes produced basis function weight significantly larger than one or smaller than zero on interpolation points. This was due to the ill-conditioned local problems, which were discussed in chapter 4. Imposing oscillatory conditions on the local flow problem bounds the values of the basis functions between zero and one and basis functions remain  $C^0$ -continuous over the element boundaries. The continuity property of basis functions on non-overlapping domains supplies unique intergrid transfer operators for the multigrid solver. The proposed multigrid solver obtains good results for this application.

## 6.2 Fluvial system

The bottom 100 ft of the SPE model represents the fluvial Upper Ness formation. The distribution of porosity and permeability was given by  $60 \times 220 \times 50$  constant cell values. Here the problem is reduced to  $60 \times 220 \times 48$  cells, to fit a hierarchy of three levels. The geological model then includes 633,600 cells and its size meets  $366 \times 671 \times 30$  m. The flow model generates 3,801,600 tetrahedral elements. The number of nodes equals 660,569. Two unknown values apply to each node. Coarsening the fine mesh produces a mesh of 79,200 cells at the first coarse level. This mesh merges to a 9,900 cell mesh on the next coarse level. The degrees of freedom reduces to 23,296 on the most coarse level. This section presents results for the top layer of the Upper Ness formation. The geological model is restricted to  $60 \times 220$  cells on the finest level. The number of elements in the hierarchy of finite element meshes is 13,200, 3,300, and 825, the number of nodes equals 13,481, 3,441, and 896.

### 6.2.1 Geological model

Figure 6.6 shows the porosity field for the fully resolved fluvial Upper-Ness formation, and Figure 6.7 presents the distribution of the porosity and the permeability for the top layer of this system. Porosity varies between 0.05 and 0.41, and permeability ranges from  $2.16 \cdot 10^{-17} \text{ m}^2$  up to  $1.9 \cdot 10^{-10} \text{ m}^2$ . For the fluvial system the permeability distribution can not be given by a single distribution. High conductivity streaks introduce heterogeneity at a larger scale than the variation in the background. The permeability distribution reveals this multiscale nature. The material behavior corresponds to the behavior given by Table 6.1 for the Tarbert heat storage problem.

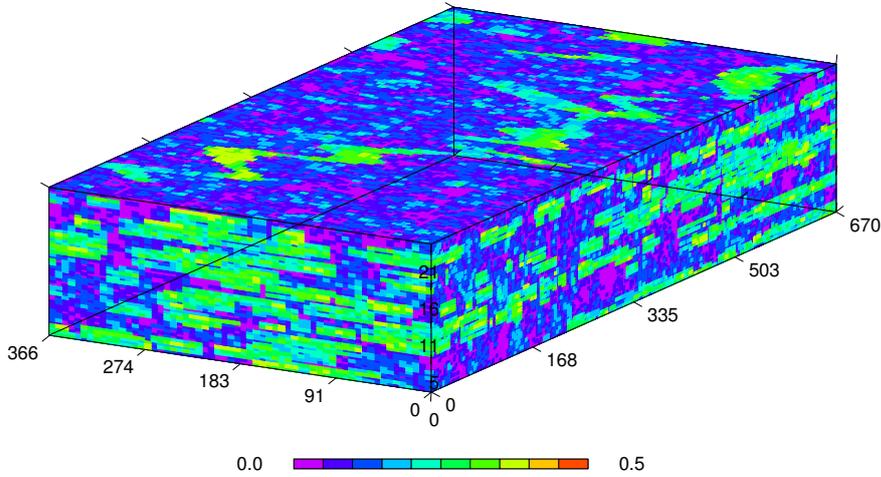


Figure 6.6: Upper-Ness porosity (-) field.

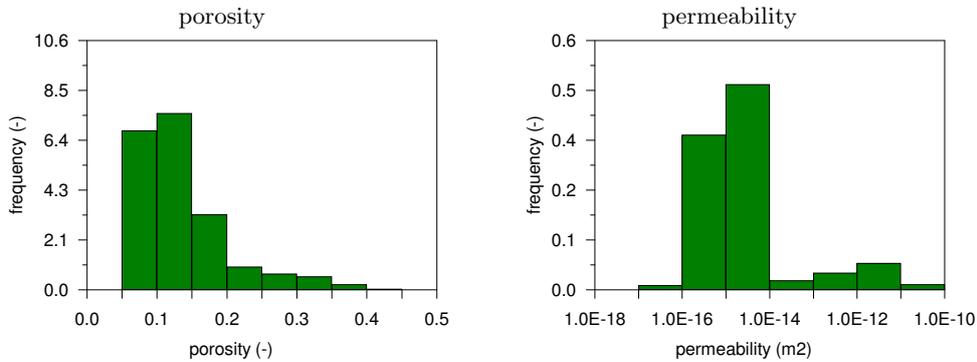


Figure 6.7: Material property distribution layer 36.

Figure 6.8 presents the porosity and permeability variation over the top layer. Fine scale parameter values are given on the third scale and upscaled values are given on

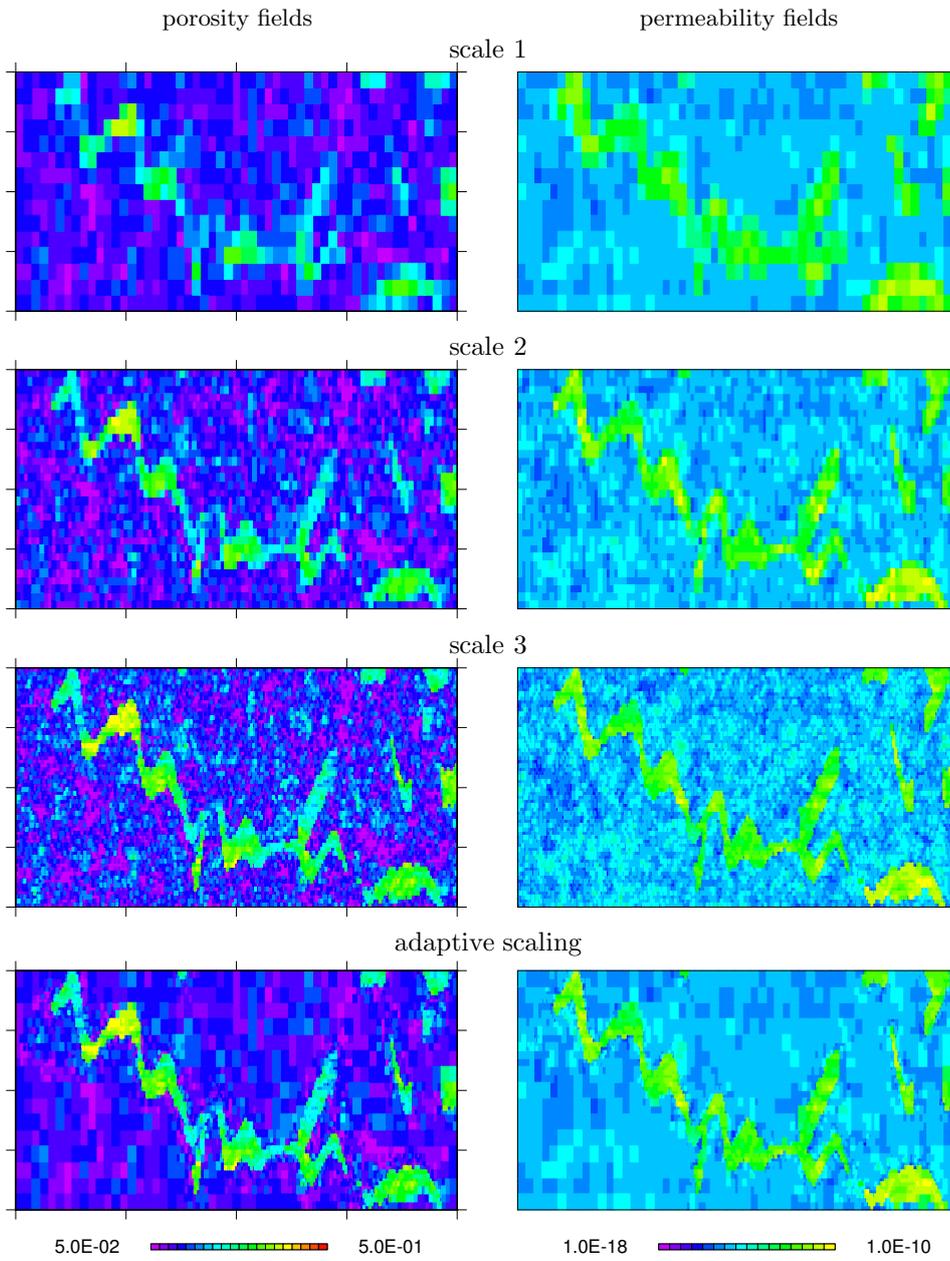


Figure 6.8: Porosity (-) and permeability ( $\text{m}^2$ ) fields, thermal energy system.

the second scale and first scale. Fixed uniform scaling shows that at coarser scales the high permeability streak is interrupted at a number of locations. Adaptive scaling captures the high conductivity streak.

### 6.2.2 Flow model

The flow model simulates a simplified geothermal energy production system. The producer well extracts hot water from the system, which can be used for heating of green houses for instance. The source well injects cooled down water into the subsurface.

The simulation considers the top layer of the fluvial system with a horizontal dimension of  $670 \times 366$  m. The natural state of the system prescribes a constant pressure of  $2 \cdot 10^7$  Pa, as the layer is positioned at a depth of 2 km, and a constant temperature of  $60^\circ$  C. The natural heat gradient was supposed to be  $3^\circ$  C per 100 m. The simulation introduces an injector well at horizontal coordinates 84 - 336 m, and a producer well at the coordinates 588 - 48 m. The distance between producer well and injector well is 580 m, which is relatively small. For this setup the cold front can be expected after a relatively short time period in the producer well. The temperature of the produced water will be low if conduction is not able to restore the temperature in the convective front. At that state the system becomes less effective or even fails. Injection and production are simulated by constant pressure wells. Injection raises the head by 200 m, giving a bottom hole pressure of  $2.2 \cdot 10^7$  Pa, and extraction lowers the head by 200 m to a corresponding pressure of  $1.8 \cdot 10^7$  Pa. The injected water has a temperature of  $10^\circ$  C. On the horizontal sides (in plane view) of the model no-flow boundary conditions apply. On the vertical ends a prescribed pressure condition of  $2 \cdot 10^7$  Pa, holds. The fine scale mesh consists of 13,481 nodes and 26,400 elements. The total simulation time of 16 year was discretized by 1000 finite difference time steps.

Figure 6.9 shows pressure fields and velocity fields for three fixed grids and one adaptively refined grid. The pressure field at the coarse scale compares well with the fine scale (level 3) pressure field. The velocity fields follow from the derivative of the pressure and are calculated less accurately for this reason. Velocity vectors decrease in size at the coarse scale because of low permeable disturbance in the high permeable streak. The proposed refinement indicator detects adjacent elements for which the normal flux on shared element edges show a large discontinuity, and exceeds the refinement criterion. Adaptivity corrects this result locally and restores the preferential flow pattern.

Figure 6.10 show the evolution of the cold front. The simulation shows that the system will be operational for more then 16 years despite the short distance between injector and producer well. However, production rates are very low due to the heterogeneous structure of the subsurface. The production rate for the layer under investigation is  $0.07 \text{ m}^3/\text{h}$  per meter filter length. In practice a filter of 60 meter is often used, which intersects many of these layers. Thermal energy systems become economically feasible for production rates of 100 to  $150 \text{ m}^3/\text{h}$  and operational periods of 90 years. Figure 6.10 compares temperature fields obtained by a computation on the fine (level 3) scale to fields on the adaptively refined mesh for a number of time steps. The results compare well.

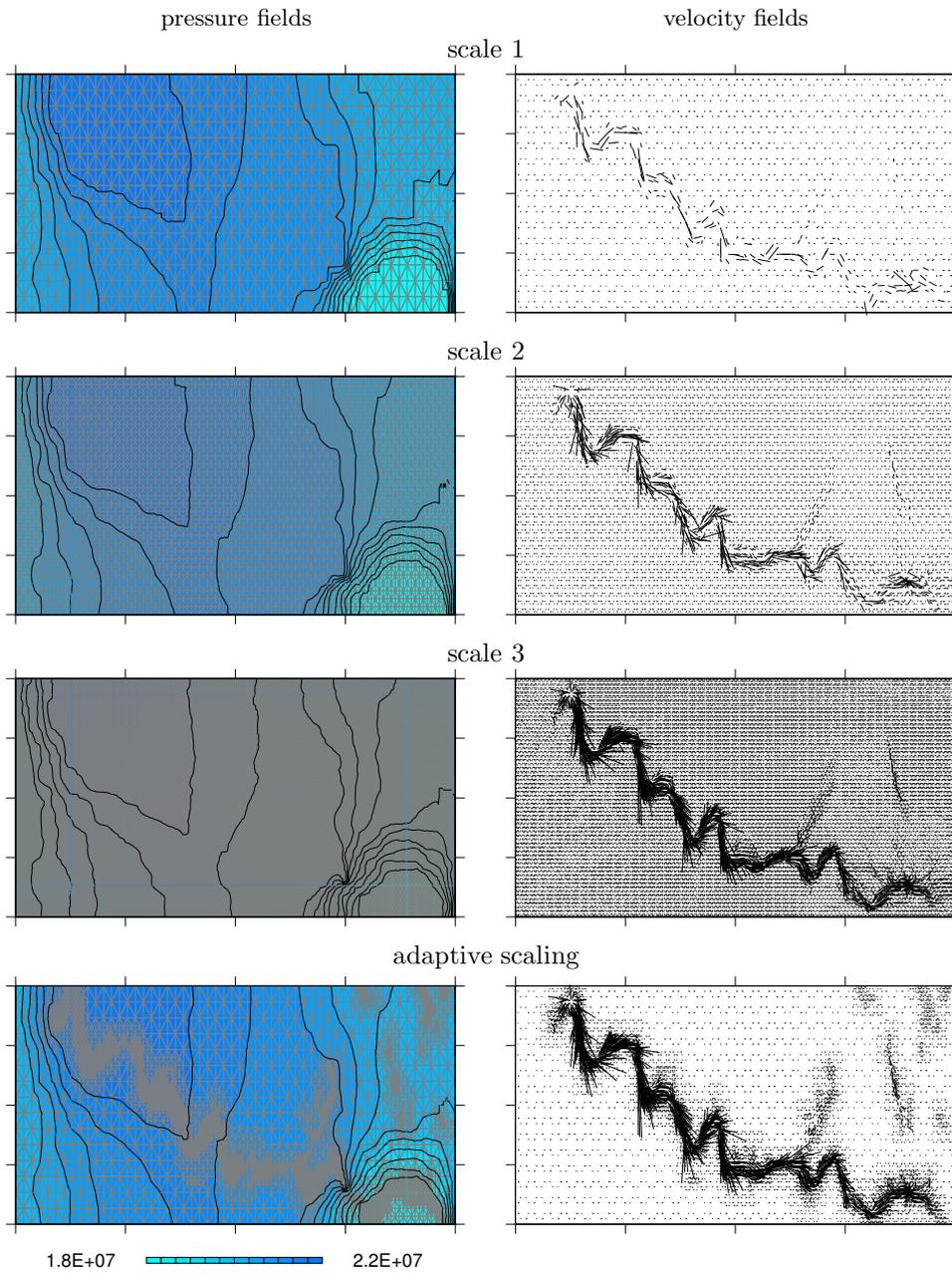


Figure 6.9: Pressure (Pa) and velocity fields, thermal energy system.

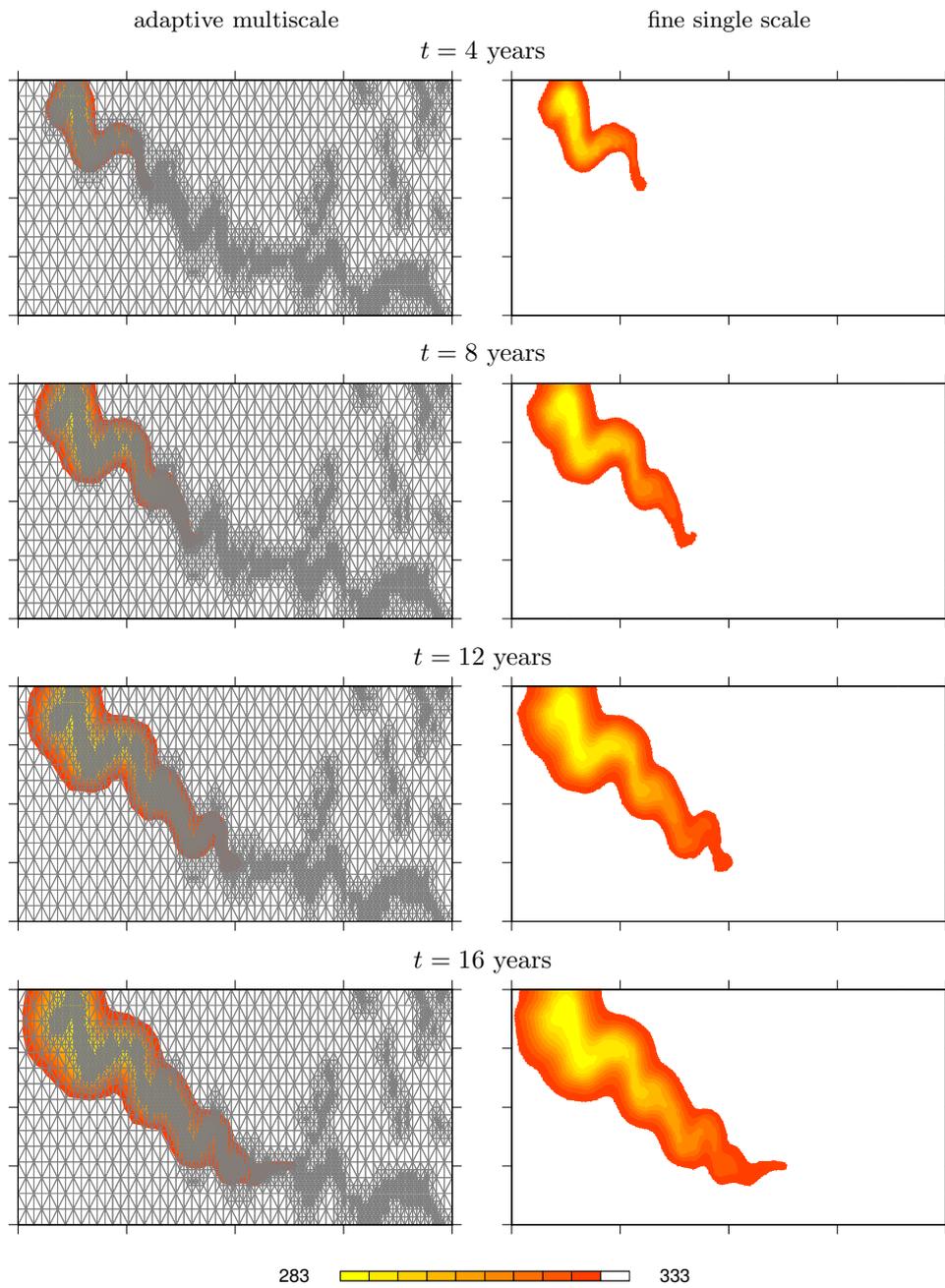


Figure 6.10: Temperature (K) fields, thermal energy system.

The coupled flow and transport problem was solved by the fully coupled system of equations presented in chapter 3 although non-linear effects are small. This approach requires setting up a system of algebraic equations for each non-linear iteration, which apply to all time steps, and solving the linearized sets of equations. The adapted mesh consists of 3,912 nodes and 6,882 elements. The size of the sets of equations have been reduced by a factor of  $3.45^2$ , the number of coefficients in the sparse matrix reduce linearly by a factor of 3.45. The pressure field was resolved by the multigrid solver outlined in chapter 4. The first step requires 17 cycles with 2 Gauss Seidel pre-smoothing operations and 2 post-smoothing operations. The convergence factor equals 0.31. A single V-cycle resolves the pressure field in the next steps because boundary conditions do not change, storage is neglected and density effects are small. The temperature field was resolved by a pre-conditioned bi-conjugate gradient solver. The computational work for setting up the systems of equations has been reduced by a factor of 3.84, because the velocity field did not change during the simulation and there was no need in constructing a new data structure for the system matrix and the right hand side vector. The overall computation time for the fully refined (level 3) simulation has been reduced by the newly developed adaptive multiscale finite element method by a factor 3.26. This application shows that simulations with an adaptively refined mesh are computationally efficient. This especially holds for three dimensional problems. However, their efficiency in reducing the number of unknowns, depends on the complexity of the structure. The transport simulation shows that the proposed refinement criterion effectively captures high permeability streaks.

# Chapter 7

## Conclusions

Simulations with an adaptively refined mesh are computationally efficient, especially for solving three-dimensional problems. Adaptivity reduces the number of finite element nodes on smooth parts of the domain. Setting up the coarsened system of equations requires less operations, the equations that need to be solved require less computer memory and the equations can be solved faster. However, adaptive techniques modify the structure of the algebraic equations in time, which produces computational overhead. Flow simulations show that a heuristic threshold on the velocity change over the flow domain reduces the update frequency, and restores the computational efficiency.

In this thesis meshes are composed out of simplex elements. Without any measurements taken, adaptivity generates hanging grid nodes and non-conforming meshes. The proposed procedure efficiently restores conformality of meshes by a regularization algorithm and an element merging procedure. The mesh regularization algorithm prevents the generation of more than one hanging grid point on a non-refined element. Remaining hanging nodes are removed by merging elements, for which a sequential coarsening approach was introduced. The applied conformal finite element method obtains nodal based mass continuity, whereas mass fluxes over the internal element boundaries are discontinuous in general. The proposed refinement indicator detects adjacent elements for which the normal flux on the shared element edge shows a large discontinuity, and exceeds the refinement criterion. Adaptivity improves the accuracy of the velocity field by refining the mesh locally. A hierarchy of finite element meshes supports the adaptive formulation. Multiscale basis functions extrapolate the coarse scale results over this hierarchy of scales. The refinement indicator operates on the finest scale. The proposed refinement criterion effectively captures complex flow patterns like high permeability streaks.

The proposed multiscale formulation captures the coarse scale behavior by modified basis functions. The functions follow from the solution of local flow problems over patches of simplex elements. Only if the patch aligns with the small scale structure, its boundaries coincide with the principal directions of the equivalent permeability tensor on the patch. Single simplex elements do not meet this condition.

A number of periodic test problems show that oversampling the patch domain obtains best results for the computation of the equivalent coarse scale behavior. As

a consequence however, the multiscale basis functions are no longer continuous over the domain boundaries and their span increases, if these basis functions are applied to obtain the coarse grid operator. Simulations on a geostatistically generated field show that the method sometimes even produces basis function weights substantially larger than one or smaller than zero on interpolation points. Sampling directly over the patch of simplex element, and closure by conditions that follow from a dimensionally reduced flow problems, enforces the compatibility conditions. For this reason closure of the local flow problems by dimensionally reduced flow problems is preferred. A second refinement criterion compares oversampled and non-oversampled function values.

Using a patch of elements can already be seen as oversampling, as the node in the center of the patch is connected to all coarse grid nodes. Pressure-dissipation averaging approximates the multiscale coarse-mesh operator. The solutions of the local flow problems follow as combinations of multiscale basis functions. The explicit computation of the equivalent permeability tensor reduces the number of integration points needed to compute the coarse scale operator in time. The procedure does not introduce new connectivities on the coarse scale. Moreover the procedure supports parallel computing.

The computation of a full equivalent permeability tensor reduces the process dependency of the variable. However, the equivalent permeability is a non-unique parameter and depends on the boundary conditions posed on its local domain. Standard multiscale finite element methods interpolate coarse scale results obtained on the global domain to the fine scale, and can not use these global results on the boundaries of the local domains. Domain decomposition techniques correct the solution of the local problem iteratively. Alternatively local-global scaling updates the equivalent permeability on subdomains with updated global information. In this thesis local refinement of the mesh corrects the pre-scaled results. The accuracy of the equivalent permeability is improved by refining the local mesh and relating the equivalent behavior on all scales to the finest grid permeability directly.

Applications of the newly developed adaptive multiscale finite element method show more accurate solutions of transport problems posed on highly heterogeneous and channelized system than solutions found on uniformly coarsened grids when compared to fine grid results.

A hierarchical formulation of coarse grid operators, as proposed by standard multigrid methods, provides less accurate equivalent material behavior compared to a direct multiscale approach. The sequential closure of local flow problems by directional lumping causes these inaccuracies. The proposed multigrid solver obtains good results, in the sense of operation counts, for the simplified heat storage and energy production applications. The intergrid transfer operators follow from multiscale weight functions obtained on non-overlapping subdomains. The computation of the coarse grid operator applies the pressure-dissipation averaging result also found on non-overlapping subdomains.

# Appendix A

## Continuum mechanics

Continuum theory aims to describe relationships between gross phenomena and averages the structure of the material at a smaller scale [85, 108]. The material of a continuum is supposed to be indefinitely divisible. Continuum mechanics applies to multiscale systems. For multiphase systems the method of volume averaging can be used to derive continuum equations [67, 123]. The technique applies spatial smoothing to translate microscopic equations, which are valid within a particular phase, to macroscopic equations, which are valid throughout the continuum. Homogenization theory assumes that macroscopic equations remain valid on the mesoscale and produces equivalent material behavior on this scale. In a porous medium, Navier-Stokes transport equations, obtained by continuum theory on the microscale, are valid in the pores. The equations are subjected to boundary conditions at the solid-liquid interface. The method of volume averaging gives a local volume averaged transport equation which corresponds to Darcy's law, if inertial effects are disregarded. Homogenization theory assumes Darcy's law to remain valid on the mesoscale and computes an equivalent permeability on this scale.

This appendix discusses the concept of continuum mechanics. Tensor calculus first presents a coordinate free notation. The microscopic formulation for transport problems presents microscopic balance equations. The macroscopic formulation uses volume averaging and translates the equations to the macroscopic scale. The mesoscopic formulation applies homogenization [129] and gives three requirements for upscaling. These requirements produce three approaches for upscaling.

### Tensor calculus

Tensor analysis will be used for the mathematical description of the mechanical and thermal behavior of the porous medium continuum. It applies a coordinate free vector and tensor notation.

The lowest order tensor is a zero order tensor or scalar. A first order tensor or vector has a length and a direction along a working line. Two vectors are identical if they have parallel working lines, are pointing in the same direction and have the same length. The vector  $\mathbf{u}$  can be written as a unique combination of components related

to three independent base vectors  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ . In index notation, the vector can then be expressed as  $\mathbf{u} = u_i \mathbf{e}_i$ . In matrix-vector notation the vector reads  $\mathbf{u} = [u]^T [\mathbf{e}]$ , or alternatively  $\mathbf{u} = [\mathbf{e}]^T [u]$ . The components of the vector follow from the inverse expression as dot-product of the vector and the coordinate base vectors. These components read  $u_i = \mathbf{u} \cdot \mathbf{e}_i$ . In matrix-vector notation the column of vector components is given by  $[u] = \mathbf{u} \cdot [\mathbf{e}]$ , or  $[u] = [\mathbf{e}] \cdot \mathbf{u}$ . The scalar product, dot-product or in-product of two vectors given by  $\mathbf{u} \cdot \mathbf{v} = u_i v_j \mathbf{e}_i \cdot \mathbf{e}_j$  reads

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|_2 \|\mathbf{v}\|_2 \cos \vartheta, \quad (\text{A.1})$$

where  $\vartheta$  is the enclosed smallest angle and  $\|\cdot\|_2$  denotes the  $L_2$ -norm which gives the length of a vector. With respect to a Cartesian vector base  $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$  and  $\mathbf{u} \cdot \mathbf{v} = u_i v_i$ . Here  $\delta_{ij}$  is the Kronecker delta function:  $\delta_{ij} = 1$  for  $i = j$  and  $\delta_{ij} = 0$  for  $i \neq j$ .

The dyadic-product, tensor-product or open product operator of two vectors given by  $\mathbf{u}\mathbf{v} = u_i v_j \mathbf{e}_i \mathbf{e}_j$  provides a linear mapping. The mapping of a vector  $\mathbf{w}$  onto  $\mathbf{r}$  is given by

$$\mathbf{u}\mathbf{v} \cdot \mathbf{w} = \mathbf{r}. \quad (\text{A.2})$$

The dyad  $\mathbf{v}\mathbf{u}$  forms the conjugate or transpose of  $\mathbf{u}\mathbf{v}$  and will be written as  $(\mathbf{u}\mathbf{v})^T$ . A second-order tensor also transforms any vector into another vector. This transformation or linear mapping is written as

$$\mathbf{A} \cdot \mathbf{u} = \mathbf{v}. \quad (\text{A.3})$$

With respect to a vector base of three independent vectors  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ , the tensor is given in index notation as  $\mathbf{A} = A_{ij} \mathbf{e}_i \mathbf{e}_j$ . In three spatial dimensions it expresses a sum of 9 dyadic-products. The matrix-vector notation reads  $\mathbf{A} = [\mathbf{e}]^T [A] [\mathbf{e}]$ . Here  $[\mathbf{e}]$  is a column of vectors and  $[A]$  is the matrix containing the tensor coefficients. The inverse relation gives the matrix of the tensor with respect to the vector base as  $A_{ij} = \mathbf{e}_i \cdot \mathbf{A} \cdot \mathbf{e}_j$ . In matrix-vector notation this expression reads  $[A] = [\mathbf{e}] \cdot \mathbf{A} \cdot [\mathbf{e}]^T$ . In general the dot-product can be written as  $\mathbf{A} \cdot \mathbf{u} = \mathbf{e}_i A_{ij} \mathbf{e}_j \cdot \mathbf{e}_k u_k$ . For a Cartesian vector base  $\mathbf{A} \cdot \mathbf{u} = A_{ij} u_j \mathbf{e}_i$  since  $\mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk}$ . A symmetric tensor is specified by  $\mathbf{A} = \mathbf{A}^T$ . With respect to an orthonormal vector base, its components satisfy  $A_{ij} = A_{ji}$ . The components of an antisymmetric tensor  $\mathbf{A} = -\mathbf{A}^T$ , then satisfy  $A_{ij} = -A_{ji}$ .

The dot-product of two tensors  $\mathbf{A} \cdot \mathbf{B} = \mathbf{e}_i A_{ij} \mathbf{e}_j \cdot \mathbf{e}_k B_{kl} \mathbf{e}_l$  follows from

$$(\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{u} = \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{u}). \quad (\text{A.4})$$

This expression should hold for all  $\mathbf{u}$ . For a Cartesian vector base the dot-product of two tensors reads  $\mathbf{A} \cdot \mathbf{B} = \mathbf{e}_i A_{ik} B_{kj} \mathbf{e}_j$ . The scalar product or double dot-product of two tensors will be written as

$$\mathbf{A} : \mathbf{B} = \text{tr}(\mathbf{A} \cdot \mathbf{B}), \quad (\text{A.5})$$

where  $\text{tr}(\cdot)$  gives the first invariant or trace of a tensor specified as  $\text{tr}(\mathbf{A}) = \mathbf{e}_i \cdot \mathbf{A} \cdot \mathbf{e}_i = A_{ii}$ . The scalar product of two tensors for a Cartesian vector base reads  $\mathbf{A} : \mathbf{B} = A_{ij} B_{ji}$ .

The inverse of a regular tensor  $\mathbf{A}$  will be written as  $\mathbf{A}^{-1}$ . If  $\det(\mathbf{A}) \neq 0$ , then

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}, \quad (\text{A.6})$$

where  $\mathbf{I} = \delta_{ij}\mathbf{e}_i\mathbf{e}_j$  denotes the unit tensor. For an orthogonal tensor  $\mathbf{A}^T \cdot \mathbf{A} = \mathbf{I}$  holds. The deviatoric part of  $\mathbf{A}$  will be indicated by  $\mathbf{A}^D$ , which is given as

$$\mathbf{A}^D = \mathbf{A} - \frac{1}{3}\text{tr}(\mathbf{A})\mathbf{I}. \quad (\text{A.7})$$

The gradient operator  $\nabla$  will be introduced for a differentiable quantity  $\mathcal{A}$  as

$$d\mathcal{A} = d\mathbf{x} \cdot \frac{\partial \mathcal{A}}{\partial \mathbf{x}} = \mathcal{A}(\mathbf{x} + d\mathbf{x}, t) - \mathcal{A}(\mathbf{x}, t) \equiv d\mathbf{x} \cdot \nabla \mathcal{A}. \quad (\text{A.8})$$

For Cartesian coordinates the gradient operator can be written as  $\nabla = \partial/\partial x_i \mathbf{e}_i$ . Here the Laplace operator  $\nabla^2$  reads  $\partial^2/\partial x_i^2 \mathbf{e}_i$ . The gradient operator produces a vector  $\nabla a$  if it works on a scalar  $a$ , and produces a dyadic-product  $\nabla \mathbf{v}$  if it operates on a vector field  $\mathbf{v}$ . For a Cartesian vector base  $\nabla a = \partial a/\partial x_i \mathbf{e}_i$  and  $\nabla \mathbf{v} = \partial v_j/\partial x_i \mathbf{e}_i \mathbf{e}_j$ . The divergence of the velocity field  $\nabla \cdot \mathbf{v}$  is the scalar field, given by the trace of the dyadic-product as

$$\nabla \cdot \mathbf{v} = \text{tr}(\nabla \mathbf{v}). \quad (\text{A.9})$$

For a Cartesian vector base  $\nabla \cdot \mathbf{v} = \partial v_j/\partial x_i \mathbf{e}_i \cdot \mathbf{e}_j = \partial v_i/\partial x_i$ .

The divergence of a tensor field  $\nabla \cdot \mathbf{A}$  is a vector field, such that for any vector  $\mathbf{v}$  the following relation holds

$$(\nabla \cdot \mathbf{A}) \cdot \mathbf{v} \equiv \nabla \cdot (\mathbf{A} \cdot \mathbf{v}) - \text{tr}(\mathbf{A}^T \cdot \nabla \mathbf{v}). \quad (\text{A.10})$$

For a Cartesian vector base  $\nabla \cdot \mathbf{A} = \partial A_{ij}/\partial x_i \mathbf{e}_j$ .

The material time derivative  $D\mathcal{A}/Dt$  of a differentiable quantity  $\mathcal{A}$  expresses the rate of change of a material particle. For a small time step  $dt$  this material derivative reads

$$\frac{D\mathcal{A}}{Dt} dt = \mathcal{A}(\mathbf{x}_0, t + dt) - \mathcal{A}(\mathbf{x}_0, t). \quad (\text{A.11})$$

Here the quantity  $\mathcal{A}$  can be a scalar field, a vector field or a tensor field. The path line or motion of every particle is given by  $\mathbf{x} = \mathbf{x}(\mathbf{x}_0, t)$  with  $\mathbf{x}_0 = \mathbf{x}(\mathbf{x}_0, t_0)$ , denoting the original position of the material point. The displacement field is given by  $\mathbf{u} = \mathbf{x}(\mathbf{x}_0, t) - \mathbf{x}_0$ , and the differential quantity of a material point is written as  $\mathcal{A}(\mathbf{x}_0, t) = \mathcal{A}(\mathbf{x}_0(\mathbf{x}, t), t) = \mathcal{A}(\mathbf{x}, t)$ . The velocity of the particle  $\mathbf{v}$  is given by the material time derivative of the position vector  $\mathbf{v} = D\mathbf{x}/Dt$ . The spatial time derivative of a differentiable quantity  $\partial \mathcal{A}/\partial t$  reads

$$\frac{\partial \mathcal{A}}{\partial t} dt = \mathcal{A}(\mathbf{x}, t + dt) - \mathcal{A}(\mathbf{x}, t). \quad (\text{A.12})$$

The relation between the material time derivative and the spatial derivative thus reads

$$\frac{D\mathcal{A}}{Dt} = \frac{\partial \mathcal{A}}{\partial t} + \mathbf{v} \cdot \nabla \mathcal{A}, \quad (\text{A.13})$$

where  $\mathbf{v} \cdot \nabla \mathcal{A}$  is the convective derivative. For a tensor-function  $f = f(\mathbf{A})$  the time derivative of the function reads

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial \mathbf{A}} : \frac{\partial \mathbf{A}}{\partial t}. \quad (\text{A.14})$$

This expression gives the rate of change at a fixed point in space.

### Microscopic formulation

Continuum mechanics at the microscopic level assumes that physical properties like density and pressure are continuously distributed in space and there exist an infinitesimally small material point as the material is regarded as indefinitely divisible. The microscopic level considers one phase only, however a number of miscible chemical species may define the material of this phase. The mass density  $\rho^\alpha$  [ $\text{ml}^{-3}$ ] of the mixture of the  $\alpha$ -phase is given by

$$\rho^\alpha = \sum_{k=1}^{n_k^\alpha} \rho_k^\alpha, \quad (\text{A.15})$$

where  $\rho_k^\alpha$  [ $\text{ml}^{-3}$ ] is the mass of the constituent  $k$  per unit volume of phase  $\alpha$ , and  $n_k^\alpha$  [-] is the number of miscible chemical species that compose the phase. The mass density of the phase mixture in general also depends on the pressure  $p$  [ $\text{ml}^{-1}\text{t}^{-2}$ ] and temperature  $T$  [K] in the phase. The mass fraction or mass-based concentration of species  $k$  in the  $\alpha$ -phase denoted by  $\omega_k^\alpha$  [-] is given as

$$\omega_k^\alpha = \frac{\rho_k^\alpha}{\rho^\alpha}. \quad (\text{A.16})$$

Conservation of mass requires that the sum of the species mass fractions is one, or  $\sum_{k=1}^{n_k^\alpha} \omega_k^\alpha = 1$ .

Global conservation laws or global balance equations apply to a material volume for which the surface moves with the same speed as the material points. The Lagrangian or material description of the continuum is associated with global conservation laws. A spatial volume or control volume is not bounded by material points. If this volume is fixed in space with respect to a reference frame, it is called a fixed control volume or fixed partial volume and associated to the Eulerian or spatial description of the continuum. Local conservation laws can be derived from global conservation laws. Local balance equations are valid in every point of the volume.

The formulation of conservation equations will be based on Reynolds's transport theorem, which uses extensive and intensive quantities. Extensive quantities like mass  $M$ , momentum  $M\mathbf{v}$  and energy  $E$  are proportional to the amount of material under consideration. Derived intensive quantities are mass per unit of mass (unity), momentum per unit of mass (velocity  $\mathbf{v}$ ), and energy per unit of mass (specific energy  $E_e$ ). Pressure  $p$ , temperature  $T$  and density  $\rho$  are intensive quantities as well.

The transport theorem relates an extensive scalar or vector quantity  $\Psi(t)$  to the corresponding continuous and differentiable intensive quantity  $\psi(\mathbf{x}, t)$ . For a given mass  $M$  at time  $t$  the transport theorem is formulated as

$$\Psi_m(t) = \int_{M(t)} \psi(\mathbf{x}, t) dM = \int_{V_m(t)} \rho(\mathbf{x}, t) \psi(\mathbf{x}, t) dV, \quad (\text{A.17})$$

where  $dM$  denotes differential mass and  $dV$  is a differential volume,  $\rho$  is the mass density, and  $V_m$  is the material volume. The rate of change of the extensive quantity

over the material volume reads

$$\frac{d\Psi_m}{dt} = \int_{V_m(t)} \left( \frac{\partial \rho \psi}{\partial t} + \nabla \cdot (\rho \psi \mathbf{v}) \right) dV = \int_{V_m(t)} \rho \frac{D\psi}{Dt} dV, \quad (\text{A.18})$$

where the results of the local equation for conservation of mass is used already. The microscopic equation for conservation of extensive mass ( $\Psi = M$ ) follows according to the transport equation for a unit intensive quantity ( $\psi = 1$ ). The local equation, which is valid for every arbitrary volume, for conservation of mass reads

$$\frac{D\rho}{Dt} + \rho \nabla \cdot \mathbf{v} = 0. \quad (\text{A.19})$$

Conservation of momentum follows from the transport equation, where the extensive quantity is momentum ( $\Psi = M\mathbf{v}$ ) and the intensive quantity is momentum per unit mass ( $\psi = \mathbf{v}$ ). The second conservation equation reads

$$\rho \frac{D\mathbf{v}}{Dt} - \nabla \cdot \boldsymbol{\sigma} - \rho \mathbf{g} = \mathbf{0}, \quad (\text{A.20})$$

where  $\boldsymbol{\sigma}$  [ $\text{ml}^{-1}\text{t}^{-2}$ ] is the (Cauchy) stress tensor and  $\mathbf{g}$  [ $\text{lt}^{-2}$ ] denotes the gravitational body force. The global equation for conservation of moment of momentum states that the rate of change of moment of momentum of a material volume is equal to the resulting moments applied by an external force. The local equation for conservation of moment of momentum simply states that the stress tensor should be equal to its transpose, formulated as  $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$ .

The first law of thermodynamics states that the rate of change of internal kinetic energy of a material volume is equal to the mechanical power performed by external loads and the supplied heat per unit of time. This first law on a microscopic level is written as

$$\rho \frac{DE_e}{Dt} - \rho r - \boldsymbol{\sigma} : \boldsymbol{\Lambda} + \nabla \mathbf{q}_e = 0, \quad (\text{A.21})$$

where  $E_e$  [-] denotes the specific internal energy,  $r$  [-] is the specific radiation density, and  $\mathbf{q}_e$  [-] denotes the heat flux vector at the surface. The rate of deformation tensor  $\boldsymbol{\Lambda}$  [ $\text{t}^{-1}$ ] denotes the symmetric part of the velocity gradient vector  $\boldsymbol{\Lambda} = \frac{1}{2} ((\nabla \mathbf{v})^T + (\nabla \mathbf{v}))$ .

As an example, the constitutive relations for a Newtonian fluid will be given and the local momentum balance equation will be derived for an incompressible fluid. For a Newtonian fluid, or linear viscous fluid the fluid stress associated with motion depends linearly on the instantaneous value of the rate of deformation.

Water and air are unable to sustain shear stresses without continuous deformation (notion of fluidity). Incompressible fluids like water undergo neglectable density changes under a wide range of loads, whereas compressible fluids like air show large density changes. For fluid substances the total stress tensor is decomposed in a viscous stress part and a pressure part. The constitutive relation between the stress tensor, pressure and rate of deformation reads

$$\boldsymbol{\sigma} = -p\mathbf{I} + \lambda \text{tr}(\boldsymbol{\Lambda})\mathbf{I} + 2\mu\boldsymbol{\Lambda}. \quad (\text{A.22})$$

Here  $\mu$  is called the first coefficient of viscosity or simply viscosity,  $\lambda + \frac{2}{3}\mu$  relates the viscous mean normal stress to the rate of change of volume and is known as the coefficient of bulk viscosity. Using this relation in the momentum balance equation (A.20) gives the Navier-Stokes equation for an incompressible fluid where  $\text{tr}\mathbf{\Lambda} = 0$ . This relation reads

$$\rho \frac{\partial \mathbf{v}}{\partial t} = \rho \mathbf{g} - \nabla p - \rho \mathbf{v} \cdot (\nabla \mathbf{v}) + \mu \nabla \cdot \nabla \mathbf{v}. \quad (\text{A.23})$$

For flow through (unfractured) porous media the acceleration term  $\rho \mathbf{v} \cdot (\nabla \mathbf{v})$  generally vanishes as velocities are relatively small.

### Macroscopic formulation

This subsection deals with continuum mechanics of the multi-phase porous medium system. The porous medium under investigation consists of a solid phase (the soil skeleton) and two immiscible fluid phases: a liquid phase (water) and a gas phase (air). Porous medium phases will be denoted by the superscript  $\alpha$ . For the multi-phase system, the solid structure is indicated by  $\alpha = s$ , the liquid phase is identified by  $\alpha = l$  and the gas phase is denoted by  $\alpha = g$ . Each phase consists of miscible chemical components called phase species. A species will be indicated by a subscript  $k$ . Species that pass through a phase interface are regarded as separate, phase-pertinent constituents. According to this definition, the total number of chemical species in the system  $n_k$  equals the sum of the species per phase  $n_k^\alpha$  and is written as

$$n_k = \sum_{\alpha=1}^{n^\alpha} n_k^\alpha. \quad (\text{A.24})$$

At the macroscopic scale the phases are described as overlapping continua. An averaging technique produces the macroscopic equations and uses generalized functions to identify the portions of space. An intensive quantity of a constituent of the liquid phase,  $\psi_k^l$  for instance, is multiplied by a generalized function which has a value one in the liquid phase and a value zero in the other phases. Leijnse [88] introduces three averaging operators for an intensive quantity: the volume averaging operator  $\langle \psi_k \rangle_v$ , the intrinsic volume averaging operator  $\langle \psi_k \rangle_v^\alpha$  and the intrinsic mass averaging operator  $\langle \psi_k \rangle_m^\alpha$ . The definition of the volume averaging operator reads

$$\langle \psi_k \rangle_v = \frac{1}{\delta V} \int_{\delta V} \psi_k dV, \quad (\text{A.25})$$

where  $\delta V$  is the volume of the representative elementary volume. The intrinsic volume average operator averages the quantity over the  $\alpha$ -phase only, as

$$\langle \psi_k \rangle_v^\alpha = \frac{1}{\delta V^\alpha} \int_{\delta V^\alpha} \psi_k dV, \quad (\text{A.26})$$

Here  $\delta V^\alpha$  denotes the volume of the  $\alpha$ -phase within the representative elementary volume. The intrinsic mass averaging operator weights the intensive quantity by the

density of the  $\alpha$ -phase and reads

$$\langle \psi_k \rangle_m^\alpha = \frac{1}{\langle \rho_k \rangle_v^\alpha \delta V^\alpha} \int_{\delta V^\alpha} \rho_k \psi_k dV. \quad (\text{A.27})$$

The relation between a volume averaged quantity and an intrinsic volume averaged quantity for a solid-fluid system is written as

$$\langle \psi_k \rangle_v = \frac{\delta V^s}{\delta V} \langle \psi_k \rangle_v^s + \frac{\delta V^f}{\delta V} \langle \psi_k \rangle_v^f. \quad (\text{A.28})$$

If the constituent exists in the fluid phase only, then  $\langle \psi_k \rangle_v = n \langle \psi_k \rangle_v^f$ , where the volume fraction  $n$  [-] denotes (volumetric) porosity. In general the volume fraction  $\phi^\alpha$  [-] of an  $\alpha$ -phase can be expressed as

$$\phi^\alpha = \frac{\delta V^\alpha}{\delta V}, \quad (\text{A.29})$$

where  $\sum_\alpha \phi^\alpha = 1$ . For a solid-fluid system the following relation holds

$$\phi^s = 1 - n, \quad \phi^f = n. \quad (\text{A.30})$$

If the liquid phase and the gas phase form the fluid phase, then the volume fraction of the fluid phase reads  $\phi^f = \phi^l + \phi^g$ . The part of the pores that is filled with the liquid phase is denoted by the saturation  $S^l$ , and  $S^g$  specifies the part of the pores filled with the gas phase. For the saturation the following relations hold  $S^l + S^g = 1$ ,  $0 \leq S^l \leq 1$ ,  $0 \leq S^g \leq 1$ . The volume fraction of the liquid phase and the gas phase can now be written as

$$\phi^l = nS, \quad \phi^g = n(1 - S), \quad (\text{A.31})$$

where  $S = S^l$  [-] denotes the degree of saturation, which equals the volume of void space occupied by the liquid relative to the volume of the void space.

The reformulated Eulerian description of the microscopic mass balance equation (A.19) for a species in the fluid phase reads

$$\frac{\partial \rho_k^f}{\partial t} + \nabla \cdot (\rho_k^f \mathbf{v}_k^f) = 0. \quad (\text{A.32})$$

The macroscopic mass balance equation follows from the microscopic mass balance equation as

$$\frac{\partial}{\partial t} (\phi^f \langle \rho_k \rangle_v^f) + \nabla \cdot (\phi^f \langle \rho_k \rangle_v^f \langle \mathbf{v}_k \rangle_m^f) + \frac{1}{\delta V} \int_\Gamma (\rho_k (\mathbf{v}_k^f - \mathbf{v}^{fs}) \cdot \mathbf{n}^{fs}) = 0, \quad (\text{A.33})$$

where  $\Gamma$  denotes the boundary of the elementary volume  $\delta V$ ,  $\mathbf{v}^{fs}$  denotes the velocity of the fluid-solid interface, and  $\mathbf{n}^{fs}$  is the outward pointing normal vector. The microscopic mass density of a phase is formed by the sum of its constituent densities as  $\rho^\alpha = \sum_{k=1}^{n_k^\alpha} \rho_k^\alpha$ . The macroscopic mass density reads  $\langle \rho \rangle_v^\alpha = \sum_{k=1}^{n_k^\alpha} \langle \rho_k \rangle_v^\alpha$ . In the formulation of the macroscopic mass balance equation, the mass averaged velocity of a fluid was used. This velocity reads

$$\langle \mathbf{v} \rangle_m^f = \frac{\sum_{k=1}^{n_k^\alpha} \langle \rho_k \rangle_v^f \langle \mathbf{v}_k \rangle_m^f}{\sum_{k=1}^{n_k^\alpha} \langle \rho_k \rangle_v^f} = \sum_{k=1}^{n_k^\alpha} \langle \omega_k \rangle_m^f \langle \mathbf{v}_k \rangle_m^f, \quad (\text{A.34})$$

where the macroscopic mass fraction follows from  $\langle \omega_k \rangle_m^\alpha = \langle \rho_k \rangle_v^\alpha / \langle \rho \rangle_v^\alpha$ , the sum of all constituents within a phase reads  $\sum_{k=1}^{n_k^\alpha} \langle \omega_k \rangle_m^\alpha = 1$  and inversely the concentration  $c_k^\alpha \equiv \langle \rho_k \rangle_v^\alpha = \langle \omega_k \rangle_m^\alpha \langle \rho \rangle_v^\alpha$ . In contrast with the volume averaged velocity, the mass averaged velocity is a physically meaningful quantity. In the absence of mass exchange at the interface the velocity of the fluid-solid interface will be zero, and the macroscopic mass balance equation (A.33) can then be written as

$$\frac{\partial}{\partial t} (\phi^f \rho^f \omega_k^f) + \nabla \cdot (\phi^f \rho^f \omega_k^f \mathbf{v}^f) + \nabla \cdot \mathbf{j}_k^f = 0, \quad (\text{A.35})$$

where the macroscopic fluid density  $\rho^f = \langle \rho_k \rangle_v^f$  and the macroscopic mass averaged velocity of the fluid phase reads  $\mathbf{v}^f = \langle \mathbf{v} \rangle_m^f$ . The non-convective flux reads

$$\mathbf{j}_k^f = \phi^f \omega_k \rho (\mathbf{v}_k - \mathbf{v}^f). \quad (\text{A.36})$$

Here the intrinsic mass average velocities, of the constituent is written as  $\mathbf{v}_k = \langle \mathbf{v}_k \rangle_m^f$ . Instead of the fluid velocity, the bulk velocity for the fluid phase or filtration velocity  $\mathbf{q}^f$  is often used in computational subsurface flow formulations. The bulk velocity expresses the fluid velocity  $\mathbf{v}^f$  relative to the macroscopic solid phase velocity  $\mathbf{v}^s$  as

$$\mathbf{q}^f = \phi^f (\mathbf{v}^f - \mathbf{v}^s). \quad (\text{A.37})$$

### Mezoscopic formulation

The sum of the mass fraction over all species in the liquid phase equals one, and the sum of the non-convective terms equals zero, according to

$$\sum_k^{n_k^\alpha} \omega_k^\alpha = 1, \quad \sum_\alpha^{n_\alpha} \mathbf{j}_k^\alpha = 0. \quad (\text{A.38})$$

Using these expressions, the steady state macroscopic mass balance equation for the fluid reads

$$\nabla \cdot (\rho \mathbf{q}) = 0, \quad (\text{A.39})$$

where the index of the fluid has been dropped. Darcy's law gives an empirical relation for the bulk velocity on the macroscale, which reads

$$\mathbf{q} = -\frac{k_r}{\mu} \mathbf{K} (\nabla p - \rho \mathbf{g}). \quad (\text{A.40})$$

Here  $k_r$  denotes the relative permeability,  $\mu$  indicates the dynamic viscosity of the fluid, and  $\mathbf{K}$  is the intrinsic permeability of the porous medium. Darcy's law holds for the macroscale and follows from the Navier-Stokes equation on the microscale. The flow equation results from the macroscopic mass balance equation and Darcy's law as

$$\nabla \cdot \left( \frac{\rho k_r}{\mu} \mathbf{K} \cdot (\nabla p - \rho \mathbf{g}) \right) = 0. \quad (\text{A.41})$$

This equation holds on the macroscopic scale and is supposed to hold on the mesoscale as well. Homogenization techniques scale macro-scale permeabilities to mezo-scale permeabilities for the flow problem. Coarse-scale permeabilities mutually relate spatially averaged pressure, flux and dissipation. Their values are derived from the spatially averaged continuity equation, Darcy's law and the dissipation equation. Zijl and Trykozko [129] proposed three requirements for upscaling: conservation of driving force, volumetric conservation of mass continuity, and conservation of energy dissipation per unit volume. A combination of these requirements produces three approaches for upscaling: pressure-flux averaging, pressure-dissipation averaging and flux-dissipation averaging. Only for periodic porous media these approaches yield the same coarse-scale permeability.

Numerical homogenization introduces an averaging operator. The spatial volume-averaging operator  $\langle \cdot \rangle$ , reads

$$\langle f \rangle = \frac{1}{V} \int_{\Omega} f d\Omega, \quad (\text{A.42})$$

where  $V = \int_{\Omega} 1 d\Omega$  denotes the volume of subdomain  $\Omega$ .

The first homogenization requirement gives an expression for the upscaled pressure, and considers conservation of driving force by pressure differences only. This expression follows from the requirement on the averaged coarse-scale pressure that should match the volume averaged fine-scale pressure distribution as  $p^{k-1} = \langle p^k \rangle$ . The superscript  $k$  denotes the fine scale and  $k - 1$  denotes the coarse scale. As  $\nabla \langle p^k \rangle = \langle \nabla p^k \rangle$ , conservation of driving force reads

$$-\nabla p^{k-1} = \langle -\nabla p^k \rangle. \quad (\text{A.43})$$

The second requirement expresses the upscaled flux components as fine-scale flux components, by considering (volumetric) conservation of mass continuity. This requirement reads

$$\nabla \cdot \mathbf{q}^{k-1} = \langle \nabla \cdot \mathbf{q}^k \rangle. \quad (\text{A.44})$$

The second requirement states that the continuity equation for the upscaled flux equals the steady state continuity for the fine-scale flux. The upscaled flux components  $\mathbf{q}^{k-1} \cdot \mathbf{e}_i$  should match the spatially averaged flux components  $\langle \mathbf{q}^k \cdot \mathbf{e}_i \rangle$ . The averaged flux components follow from Darcy's law as  $\langle \mathbf{q}^k \cdot \mathbf{e}_i \rangle = \langle -\mathbf{K} \cdot \nabla p^k \cdot \mathbf{e}_i \rangle$ , where  $p$  equals pressure minus the hydrostatic pressure, and the mobility is assumed to be constant.

The third requirement considers conservation of energy dissipation per unit volume

$$-\nabla p^{k-1} \cdot \mathbf{q}^{k-1} = \langle -\nabla p^k \cdot \mathbf{q}^k \rangle. \quad (\text{A.45})$$

According to this condition the dissipation equation on the coarse scale matches the averaged dissipation equation on the fine scale. The dissipation equation expresses the rate of irreversible conversion of mechanical energy to internal energy within the upscaling domain (per unit volume).

In general three requirements over-specify the upscaling problem, and only two out of three requirements can be satisfied simultaneously. The first averaging approach, called pressure-flux averaging, combines the first and second requirement as

$$\langle \nabla p^{(a)k} \rangle \tilde{\mathbf{K}} = \langle \mathbf{q}^{(a)k} \rangle, \quad (\text{A.46})$$

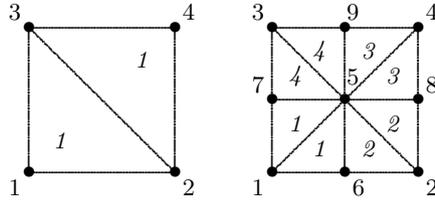


Figure A.1: Triangular parent and child element patches.

where  $a = (1 \dots n_d)$  indicates the loading case, and  $\tilde{\mathbf{K}}$  denotes the equivalent permeability tensor.

The second averaging approach combines the first requirement and third requirement and is named pressure-dissipation averaging. The set of algebraic equations for this approach reads

$$\langle -\nabla p^{(a)k} \rangle \cdot \tilde{\mathbf{K}} \cdot \langle -\nabla p^{(b)k} \rangle = \langle -\nabla p^{(a)k} \cdot \mathbf{q}^{(b)k} \rangle, \quad (\text{A.47})$$

here both  $(a)$  and  $(b)$  indicate loading cases.

Flux-dissipation averaging forms the third averaging approach, and combines the second and the third requirement as

$$\langle \mathbf{q}^{(a)k} \rangle \cdot \tilde{\mathbf{K}}^{-1} \cdot \langle \mathbf{q}^{(b)k} \rangle = \langle -\nabla p^{(a)k} \cdot \mathbf{q}^{(b)k} \rangle. \quad (\text{A.48})$$

All three expressions present a set of four algebraic equations in two dimensions, and give nine equations in a three-dimensional space.

Homogenization theory applies to (structurally and functionally) periodic fields. For these fields analytical expressions of the effective coarse scale permeability are available. This makes the fields convenient for comparison [59, 82, 129]. Numerical homogenization techniques approximate the effective behavior of these fields by imposing periodic boundary conditions. Periodic boundary conditions for the patch of elements depicted by figure A.1 follow from

$$\begin{aligned} p_2 &= p_1 + \Delta p^x & p_8 &= p_7 + \Delta p^x, \\ p_3 &= p_1 + \Delta p^y & p_9 &= p_6 + \Delta p^y, \\ p_4 &= p_1 + \Delta p^x + \Delta p^y. \end{aligned} \quad (\text{A.49})$$

These relations reduce the set of equations, which followed from the finite element

discretization of the steady state flow problem, as

$$\begin{bmatrix}
 a_{11} & a_{15} & a_{16} & a_{17} \\
 a_{22} & a_{25} & a_{26} & a_{28} \\
 a_{33} & a_{35} & a_{39} & a_{37} \\
 a_{44} & a_{45} & a_{49} & a_{48} \\
 a_{51} + a_{52} + a_{53} + a_{54} & a_{55} & a_{56} + a_{59} & a_{57} + a_{58} \\
 a_{61} + a_{62} & a_{65} & a_{66} & a_{67} + a_{68} \\
 a_{71} + a_{73} & a_{75} & a_{76} + a_{79} & a_{77} \\
 a_{82} + a_{84} & a_{85} & a_{86} + a_{89} & a_{88} \\
 a_{93} + a_{94} & a_{95} & a_{99} & a_{97} + a_{98}
 \end{bmatrix}
 \begin{bmatrix}
 p_1 \\
 p_5 \\
 p_6 \\
 p_7
 \end{bmatrix}
 =
 \begin{bmatrix}
 Q_1 \\
 Q_2 \\
 Q_3 \\
 Q_4 \\
 Q_5 \\
 Q_6 \\
 Q_7 \\
 Q_8 \\
 Q_9
 \end{bmatrix}
 +
 \begin{bmatrix}
 -a_{22} - a_{28} & & & \\
 & -a_{33} - a_{39} & & \\
 -a_{44} - a_{48} & -a_{44} - a_{49} & & \\
 -a_{52} - a_{54} - a_{58} & -a_{53} - a_{54} - a_{59} & & \\
 -a_{62} - a_{68} & & & \\
 & -a_{73} - a_{79} & & \\
 -a_{82} - a_{84} - a_{88} & -a_{84} - a_{89} & & \\
 -a_{94} - a_{98} & -a_{93} - a_{94} - a_{99} & &
 \end{bmatrix}
 \begin{bmatrix}
 \Delta p^x \\
 \Delta p^y
 \end{bmatrix}. \quad (\text{A.50})$$

Continuity of flow suggests

$$Q_1 + Q_2 + Q_3 + Q_4 = 0, \quad Q_6 + Q_9 = 0, \quad Q_7 + Q_8 = 0 \quad Q_5 = 0. \quad (\text{A.51})$$

Finally, these conditions reduce the set of algebraic equations to four, from which the fine scale pressure unknowns  $p_5$ ,  $p_6$ , and  $p_7$  can be calculated given a coarse scale pressure  $p_1$ . The remaining pressures result from equation (A.49).

In this section numerical homogenization by pressure-dissipation averaging will be applied for a number of periodic fields. Multiscale basis functions reconstruct the pressure fields of local flow problems. The basis functions follow either from an oversampling procedure or from closure by oscillating conditions of the subdomain. Appendix B includes the multiscale basis functions for all fields. The pressure fields will be compared with fields that follow from closure by periodic loading conditions.

Figure A.2, A.3, A.4 and A.5 show the local flow fields for the horizontally layered cell, the checker board cell, the cell with an inclusion and the diagonally layered cell. The upper row of plots show the reconstructed flow fields that follow from summing basis functions obtained by the oversampling procedure. Here the sample domain was extended by a zone with a thickness of a complete periodic cell. The middle row of plots shows a reconstruction by basis functions that follows from the oscillating closure procedure. Results from both procedures can be compared with local flow fields that follow from loading with periodic boundary conditions. These results are given by the lower plots. The effective permeability follows from homogenization theory and the equivalent permeability follows from pressure-dissipation averaging. Equivalent permeability is obtained on a number of discretization scales, and gathered in Table A.1, A.2, A.3, and A.4.

Table A.1: Equivalent permeability for the horizontally layered cell.

nodes	$k_{xx}$	$k_{xy}$	$k_{yx}$	$k_{yy}$	
	$5.05 \cdot 10^{-1}$	0	0	$1.98 \cdot 10^{-2}$	periodic
25	$5.05 \cdot 10^{-1}$	$-9.71 \cdot 10^{-16}$	$-1.09 \cdot 10^{-15}$	$1.98 \cdot 10^{-2}$	oversampled
81	$5.05 \cdot 10^{-1}$	$-8.31 \cdot 10^{-16}$	$-1.76 \cdot 10^{-15}$	$1.98 \cdot 10^{-2}$	
25	$5.05 \cdot 10^{-1}$	$-1.11 \cdot 10^{-15}$	$-1.11 \cdot 10^{-15}$	$1.98 \cdot 10^{-2}$	oscillating
81	$5.05 \cdot 10^{-1}$	$-1.59 \cdot 10^{-15}$	$-1.53 \cdot 10^{-15}$	$1.98 \cdot 10^{-2}$	

Table A.2: Equivalent permeability for the checkerboard cell.

nodes	$k_{xx}$	$k_{xy}$	$k_{yx}$	$k_{yy}$	
	$1.00 \cdot 10^{-1}$	0	0	$1.00 \cdot 10^{-1}$	periodic
25	$4.07 \cdot 10^{-1}$	$-7.57 \cdot 10^{-4}$	$-7.57 \cdot 10^{-4}$	$3.85 \cdot 10^{-1}$	oversampled
81	$2.99 \cdot 10^{-1}$	$-3.83 \cdot 10^{-4}$	$-3.83 \cdot 10^{-4}$	$2.99 \cdot 10^{-1}$	
25	$4.34 \cdot 10^{-1}$	$-9.90 \cdot 10^{-3}$	$-9.90 \cdot 10^{-3}$	$4.34 \cdot 10^{-1}$	oscillating
81	$3.87 \cdot 10^{-1}$	$-3.12 \cdot 10^{-2}$	$-3.12 \cdot 10^{-2}$	$3.87 \cdot 10^{-1}$	

Table A.3: Equivalent permeability for the domain with single inclusion.

nodes	$k_{xx}$	$k_{xy}$	$k_{yx}$	$k_{yy}$	
	$5.92 \cdot 10^{-1}$	0	0	$5.92 \cdot 10^{-1}$	periodic
25	$6.31 \cdot 10^{-1}$	$-4.92 \cdot 10^{-16}$	$-4.36 \cdot 10^{-16}$	$6.31 \cdot 10^{-1}$	oversampled
81	$5.92 \cdot 10^{-1}$	$-7.78 \cdot 10^{-16}$	$-2.21 \cdot 10^{-15}$	$5.92 \cdot 10^{-1}$	
25	$6.42 \cdot 10^{-1}$	$-6.45 \cdot 10^{-16}$	$-6.70 \cdot 10^{-16}$	$6.42 \cdot 10^{-1}$	oscillating
81	$6.10 \cdot 10^{-1}$	$-1.05 \cdot 10^{-15}$	$-8.87 \cdot 10^{-16}$	$6.10 \cdot 10^{-1}$	

Table A.4: Equivalent permeability for the diagonally layered cell.

nodes	$k_{xx}$	$k_{xy}$	$k_{yx}$	$k_{yy}$	
	$2.62 \cdot 10^{-1}$	$2.42 \cdot 10^{-1}$	$2.42 \cdot 10^{-1}$	$2.62 \cdot 10^{-1}$	periodic
25	$2.72 \cdot 10^{-1}$	$2.34 \cdot 10^{-1}$	$2.34 \cdot 10^{-1}$	$2.72 \cdot 10^{-1}$	oversampled
81	$2.65 \cdot 10^{-1}$	$2.39 \cdot 10^{-1}$	$2.39 \cdot 10^{-1}$	$2.65 \cdot 10^{-1}$	
25	$5.82 \cdot 10^{-1}$	$5.44 \cdot 10^{-1}$	$5.44 \cdot 10^{-1}$	$5.82 \cdot 10^{-1}$	oscillating
81	$4.68 \cdot 10^{-1}$	$4.42 \cdot 10^{-1}$	$4.42 \cdot 10^{-1}$	$4.68 \cdot 10^{-1}$	

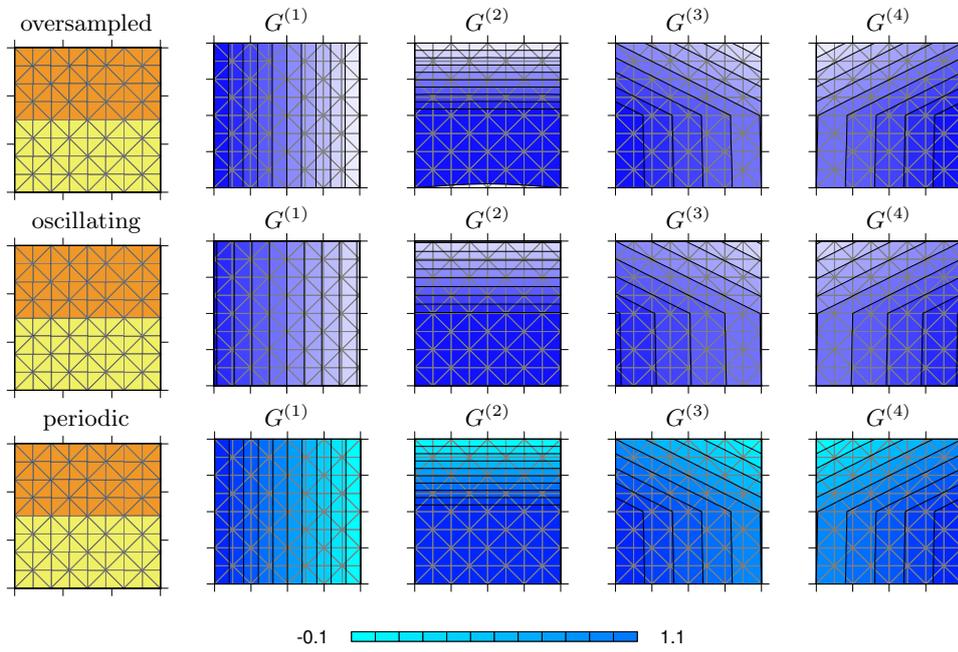


Figure A.2: Pressure distribution on a horizontally layered field.

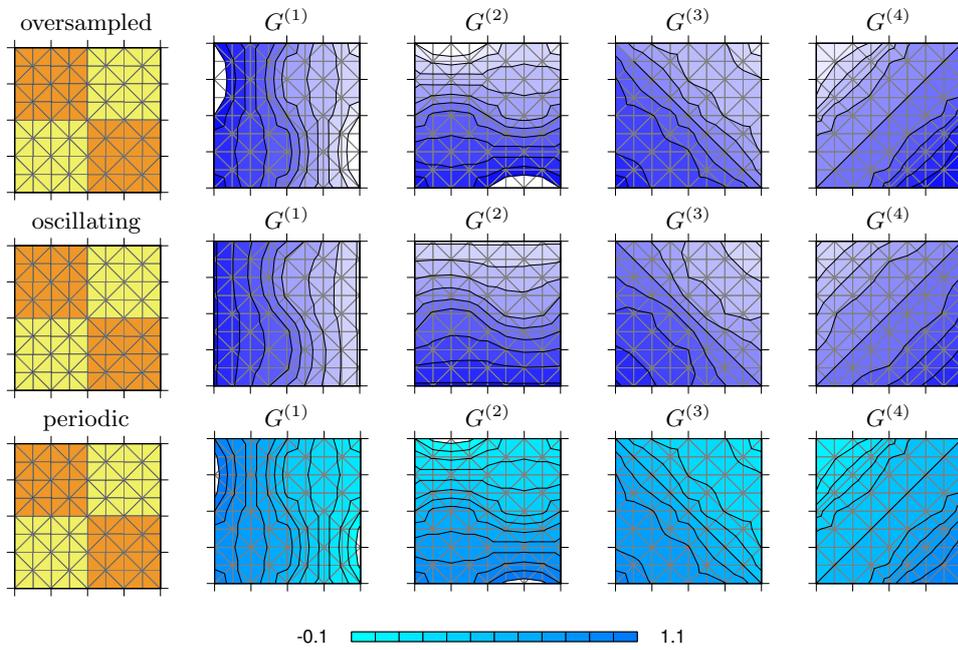


Figure A.3: Pressure distributions on a checkerboard field.

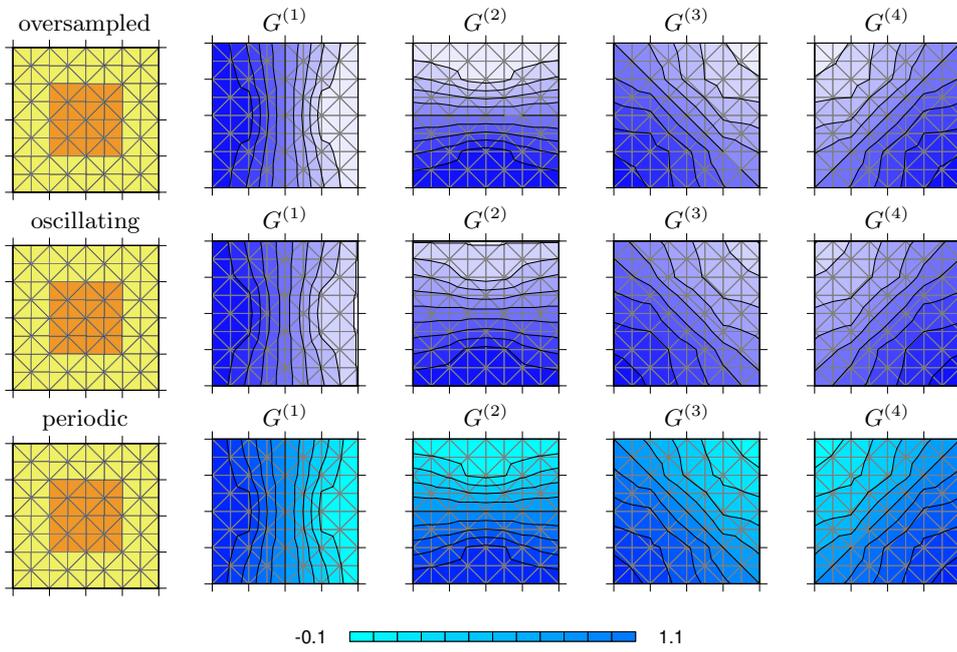


Figure A.4: Pressure distribution on a domain with inclusions.

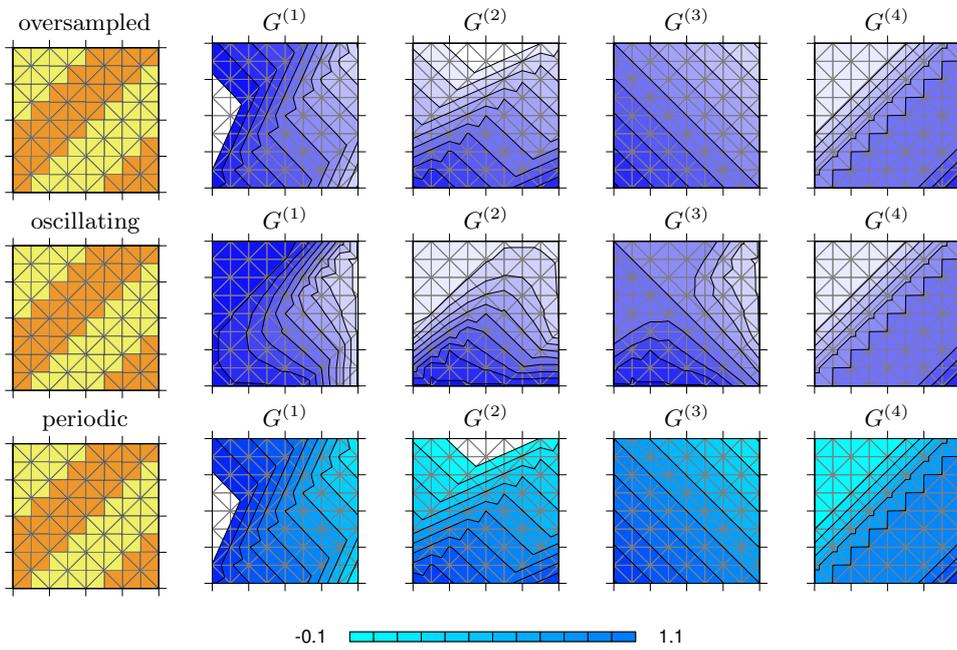


Figure A.5: Pressure distribution on a diagonally layered field.

# Appendix B

## Finite element method

The finite element method is a numerical method for solving differential equations [10, 70, 127, 128, 14]. The method applies the concept of piecewise approximations and interconnected elements in a discretization step. The elements may have simple triangular shapes or tetrahedral shapes. Elements are connected through nodes. For the case of linear elements, the nodes are located in the corners of the elements. Compatibility conditions need to be satisfied when connecting element boundaries. After discretization, local matrix expressions are developed, which relate the nodal variables of each element. Next, the element matrices are assembled to form the global set of equations. Boundary conditions are incorporated into the resulting global matrix-vector equation. Finally, the global set of equations has to be solved.

This appendix introduces the finite element method, which applies to the weak formulation for a differential equation. Numerical integration in space and integration in time are briefly discussed, followed by coordinate transformation for vector and tensor properties. Basis functions and weighting functions are given for a linear element in global coordinates. Isoparametric elements present the basis functions on local coordinates and provide a convenient mapping to global coordinates. Next, the element hierarchy will be presented for simplex elements. This hierarchy will be used for the construction of multiscale basis functions.

### Finite element formulation

A general partial differential problem introduces the finite element method. This problem can be written as

$$\mathcal{L}(u) = f \quad \text{on } \Omega, \quad (\text{B.1})$$

where  $\Omega$  is the domain, and  $u = u(\mathbf{x}, t)$  denotes the exact solution of the unknown, which can be a function of space  $\mathbf{x}$  and time  $t$ . For nonlinear problems the operator  $\mathcal{L}$  and the force vector  $f$  may be functions of space and time but may also depend on the unknown  $u$ . The finite element method applies a trial function  $\hat{u}(\mathbf{x}, t)$  that approximates the unknown function  $u$ . In this thesis the trial functions will be written as

$$\hat{u}(\mathbf{x}, t) = u_a(t)N_a(\mathbf{x}), \quad a = 1, \dots, n_n, \quad (\text{B.2})$$

where  $N_a$  are linear independent basis functions or shape functions, which do not depend on time. The number of independent basis functions is expressed by  $n_n$ . The basis functions are specified over the entire solution domain and need to satisfy the essential boundary conditions. In general, the trial function will not satisfy equation (B.1) and a residual  $r$  can be calculated as

$$r = \mathcal{L}(\hat{u}) - f. \quad (\text{B.3})$$

The method of weighted residuals calculates the trial function coefficients  $u_a$  such that the residual  $r$  is minimized in some sense. Hence, a weighted integral of the residual is formed over the entire solution domain, and a linear combination of linearly independent trial functions is considered. The weighted residual integral is then set to zero according to

$$\int_{\Omega} W_a r d\Omega = \int_{\Omega} W_a (\mathcal{L}(\hat{u}) - f) d\Omega = 0, \quad a = 1, \dots, n_n. \quad (\text{B.4})$$

The residual is orthogonalized to  $n_n$  linear independent weight functions  $W_a$ . The Galerkin weighted residual method chooses the weight functions identical to the basis functions. The basis functions are defined piecewise in each element, which restricts integration over these non-overlapping subdomains. Basis functions or shape functions are in fact interpolation functions over an element. The interpolation functions are associated with a single node of the element. The conformal finite element method specifies compatibility conditions, which require that interpolation functions need to be  $C_0$ -continuous over the element boundaries and the following condition holds

$$N_a(\mathbf{x}_b) = \delta_{ab}, \quad (\text{B.5})$$

where  $\delta_{ab}$  denotes the Dirac delta function, which has a value 1 if  $a = b$  and 0 if  $a \neq b$ . The global coordinate of node  $b$  is written as  $\mathbf{x}_b$ . The global set of algebraic equations then follows from the assembly of local sets. This process is written as

$$\int_{\Omega} W_a (\mathcal{L}(N_b u_b) - f) d\Omega = \bigg\bigg\bigg\int_{e=1}^{n_e} W_a^e (\mathcal{L}(N_b u_b) - f)^e d\Omega, \quad (\text{B.6})$$

where  $n_e$  is the number of elements, and  $\bigg\bigg\bigg\int_{e=1}^{n_e}$  denotes the assemblage operator. This expression presents the global set of simultaneous equations for  $a$ , where  $b = 1, \dots, n_n$ .

### Weak formulation

An essential part of the finite element method is the weak formulation of a mathematical problem. This weak formulation will be discussed in more detail for an elliptic test problem. Hughes [70] posed this problem as

$$\frac{\partial^2 u}{\partial x^2} = -f, \quad -\frac{\partial u(0)}{\partial x} = h, \quad u(1) = g, \quad (\text{B.7})$$

where  $f$  is a smooth scalar-valued function defined on the closed unit interval  $[0, 1]$  and  $f(x)$  is a real function in  $[0, 1]$ . The interval  $[0, 1]$  is the domain of the problem,

which will be indicated by  $\Omega$  and  $\mathfrak{R}$  is the range. The strong, or classical form of the boundary-value problem is stated as: given  $f : \Omega \rightarrow \mathfrak{R}$  and  $g$  and  $h$ , find  $u : \Omega \rightarrow \mathfrak{R}$ , such that equation (B.7) is satisfied.

At first two classes of functions need to be characterized in order to define the weak or variational form of the strong formulation. The first class is composed of trial solutions  $\hat{u}$ , which need to satisfy  $\hat{u}(1) = g$ . The derivatives of these trial solutions are required to be square integrable, so

$$\int_0^1 \left( \frac{\partial \hat{u}}{\partial x} \right)^2 d\Omega < \infty. \quad (\text{B.8})$$

$H^1$ -functions from a Hilbert-space satisfy this condition. If two functions  $\hat{u}$  and  $w$  are square integrable themselves, then their in-product exists, so

$$\int_0^1 \hat{u} w d\Omega < \infty. \quad (\text{B.9})$$

The set of trial functions  $\mathcal{U}$  now consists of all functions which have square integrable derivatives and have a value  $g$  on  $x = 1$ . This selection is written as

$$\mathcal{U} = \{ \hat{u} | \hat{u} \in H^1, \hat{u}(1) = g \}, \quad (\text{B.10})$$

where  $\hat{u}$  is a member of the set. The collection of weighting functions has a homogeneous boundary condition for  $x = 1$ . This collection  $\mathcal{W}$  will be written as

$$\mathcal{W} = \{ w | w \in H^1, w(1) = 0 \}. \quad (\text{B.11})$$

A weak formulation of the strong form given by equation (B.7) is found by multiplying both sides with a smooth (interpolation) function  $w$  and by integration over the domain. This weak form reads

$$\int_0^1 w \frac{\partial^2 u}{\partial x^2} dx = - \int_0^1 w f dx. \quad (\text{B.12})$$

Application of Green's theorem reduces the order of differentiation and gives a lower order weak form, written as

$$\int_0^1 \frac{\partial u}{\partial x} \frac{\partial w}{\partial x} dx = \int_0^1 w f dx + w(0)h. \quad (\text{B.13})$$

The weak formulation is stated as: given  $f : \Omega \rightarrow \mathfrak{R}$ ,  $g$  and  $h$ , find  $u : \Omega \rightarrow \mathfrak{R}$ , such that equation (B.13) is satisfied. This formulation is often called the virtual work equation or virtual displacement equation in mechanics. It should be noted that equation (B.12) and equation (B.13) are both weak forms of equation (B.7). However, equation (B.13) is preferred in finite element formulations because the required order of the interpolation functions is lower.

For non-homogeneous natural boundary conditions Green's theorem introduces a boundary integral. For the case of a multi-dimensional problem Green's theorem

rewrites the integral over the weighted divergence of a flux field with components  $v_i$  in two steps. Integration by parts in the first step gives

$$\int_{\Omega} w \frac{\partial v_i}{\partial x_i} d\Omega = \int_{\Omega} \frac{\partial}{\partial x_i} (w v_i) d\Omega - \int_{\Omega} v_i \frac{\partial w}{\partial x_i} d\Omega. \quad (\text{B.14})$$

The second step applies Gauss theorem on the first term in the right hand side, so that equation (B.14) becomes

$$\int_{\Omega} w \frac{\partial v_i}{\partial x_i} d\Omega = - \int_{\Omega} v_i \frac{\partial w}{\partial x_i} d\Omega + \int_{\Gamma} w v_i n_i d\Gamma, \quad (\text{B.15})$$

where  $n_i$  are the components of a vector on the boundary of the domain, pointing outside.

### Coordinate transformations

Coordinate transformations facilitate the finite element formulation of multidimensional partial differential equations. The mapping of a two-dimensional vector  $\mathbf{g}$  given in global coordinates  $(g_x, g_y)$  into its local expression  $(g_{\xi}, g_{\eta})$  can be given by

$$\begin{bmatrix} g_{\xi} \\ g_{\eta} \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} g_x \\ g_y \end{bmatrix}, \quad (\text{B.16})$$

where  $\alpha$  is known as an Eulerian angle. According to figure B.1  $\cos \alpha = dx/d\xi = dy/d\eta$  and  $\sin \alpha = dy/d\xi = -dx/d\eta$ . More general, the Jacobian matrix transforms the

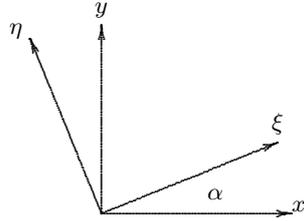


Figure B.1: Coordinate transformation.

coordinates from the global system  $(x, y)$  to the local system  $(\xi, \eta)$  as

$$[J] \equiv \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} \\ \frac{\partial x}{\partial \eta} & \frac{\partial y}{\partial \eta} \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}. \quad (\text{B.17})$$

The back transformation is achieved by the inverse of the Jacobian matrix. This inverse matrix follows from

$$[J]^{-1} \equiv \begin{bmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \eta}{\partial x} \\ \frac{\partial \xi}{\partial y} & \frac{\partial \eta}{\partial y} \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}. \quad (\text{B.18})$$

By definition the matrix product of both gives the unit matrix. The vector transformation from the local system  $(\xi, \eta)$  to the global system  $(x, y)$  and backward reads

$$[g^\xi] = [J^{-1}]^T [g^x], \quad [g^x] = [J]^T [g^\xi]. \quad (\text{B.19})$$

In index notations the relationship between the vector components in global coordinates  $g_i^x$  into its local coordinate components  $g_i^\xi$  is expressed by

$$g_i^\xi = g_j^x \frac{\partial \xi_i}{\partial x_j}, \quad g_i^x = g_j^\xi \frac{\partial x_i}{\partial \xi_j}. \quad (\text{B.20})$$

The conductivity tensor  $\mathbf{K}$  has orthotropic properties. The principal conductivities are given along their principal directions. These directions are turned against the global Cartesian coordinate system by rigid body rotation. If the conductivity matrix is a diagonal matrix, then its components contain the conductivities along the principal directions, and a rotation matrix maps this matrix into global Cartesian coordinates. The transformation between coordinates  $(x, y)$  and  $(\xi, \eta)$  may be written as

$$\begin{bmatrix} K_{\xi\eta} & K_{\xi\eta} \\ K_{\eta\xi} & K_{\eta\eta} \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} K_{xx} & K_{xy} \\ K_{yx} & K_{yy} \end{bmatrix} \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}, \quad (\text{B.21})$$

or more general by

$$[K^\xi] = [J^{-1}]^T [K^x] [J^{-1}], \quad [K^x] = [J]^T [K^\xi] [J]. \quad (\text{B.22})$$

In index notations the relations between the tensor components in local rotated coordinates  $K_{ij}^\xi$  and global coordinates  $K_{ij}^x$  read

$$K_{ij}^\xi = K_{kl}^x \frac{\partial \xi_i}{\partial x_k} \frac{\partial \xi_j}{\partial x_l}, \quad K_{ij}^x = K_{kl}^\xi \frac{\partial x_i}{\partial \xi_k} \frac{\partial x_j}{\partial \xi_l}. \quad (\text{B.23})$$

Two-dimensional anisotropy rotates the principal axes into the global Cartesian coordinates by a given angle  $\alpha$  between the first principal axis of the permeability tensor and the first Cartesian direction.

### Isoparametric elements

Isoparametric elements support a convenient method to obtain the integrals over the elements by mapping the global element to a local element. Figure B.2 illustrates the mapping of a triangular element from global to local coordinates. The transformation

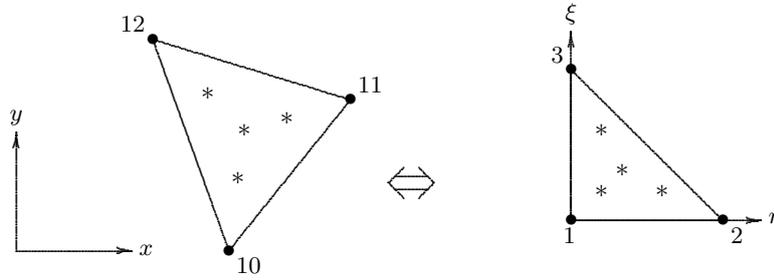


Figure B.2: Triangular elements, global and local coordinates.

from local coordinates  $\xi$  of an isoparametric element to global coordinates  $\mathbf{x}$  with vector component  $x_i$  of the actual element is obtained by  $x_i(\xi) = x_{ai}N_a(\xi)$ .  $N_a(\xi)$  denotes the basis function value at location  $\xi$  and subscript  $a$  is the node number. This expression will be written for short as

$$x_i = x_{ai}N_a. \quad (\text{B.24})$$

As an example, the shape functions for a triangular element with local nodal coordinates  $(\xi, \eta)_1 = (0, 0)$ ,  $(\xi, \eta)_2 = (1, 0)$  and  $(\xi, \eta)_3 = (0, 1)$  read

$$N_1 = 1 - \xi - \eta, \quad N_2 = \xi, \quad N_3 = \eta. \quad (\text{B.25})$$

According to equation (B.24) the global coordinates  $(x, y)$  of a point inside the local element with local coordinates  $(\xi, \eta)$  follow explicitly from

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} + \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}. \quad (\text{B.26})$$

The inverse form obtains the local coordinates from the global coordinates. For this mapping a set of equations needs to be solved.

To obtain global derivatives of the shape functions, the Jacobian matrix and its inverse have to be calculated. The components  $J_{ij}$  of the Jacobian matrix follow from

$$J_{ij} \equiv \frac{\partial x_j}{\partial \xi_i} = \frac{\partial N_a}{\partial \xi_i} x_{aj}. \quad (\text{B.27})$$

For a linear triangular element the Jacobian matrix is invariant over the element. This is for example not the case for a bi-linear quadrilateral element. Components of the inverse of the Jacobian matrix are expressed as

$$J_{ij}^{-1} \equiv \frac{\partial \xi_j}{\partial x_i}. \quad (\text{B.28})$$

By definition  $J_{ik}J_{kj}^{-1} = \delta_{ij}$  where  $\delta_{ij}$  is the Dirac delta. For a one-dimensional element the Jacobian is a scalar  $J$  and its inverse is written as  $1/J$ . In multiple dimensions the

Jacobian inverse can be found by Cramer's rule. The global derivatives of the shape function now read

$$\frac{\partial N_a}{\partial x_i} = \frac{\partial \xi_j}{\partial x_i} \frac{\partial N_a}{\partial \xi_j} = J_{ij}^{-1} \frac{\partial N_a}{\partial \xi_j}. \quad (\text{B.29})$$

For spatial integration the absolute value of the determinant of the Jacobian matrix  $|J| \equiv |\det [J]|$  is needed. For a two-dimensional case the Jacobian determinant reads

$$|J| = \left| \frac{\partial x_1}{\partial \xi_1} \frac{\partial x_2}{\partial \xi_2} - \frac{\partial x_2}{\partial \xi_1} \frac{\partial x_1}{\partial \xi_2} \right|. \quad (\text{B.30})$$

### Basis functions

This subsection introduces basis functions on a single level. Figure B.3 presents a linear and a quadratic line element and their nodal locations. The local shape functions for



Figure B.3: Line element, linear and quadratic.

a line element with local nodal coordinates  $(\xi)_1 = (0)$  and  $(\xi)_2 = (1)$  are given by

$$N_1 = 1 - \xi, \quad N_2 = \xi. \quad (\text{B.31})$$

For a quadratic element with nodal coordinates  $(\xi)_1 = (0.0)$ ,  $(\xi)_2 = (1.0)$  and  $(\xi)_3 = (0.5)$ , the shape functions read

$$N_1 = 2\xi^2 - 3\xi + 1, \quad N_2 = 2\xi^2 - \xi, \quad N_3 = -4\xi^2 + 4\xi. \quad (\text{B.32})$$

The left picture of figure B.4 shows a linear three-node triangular element and the picture in the right displays a six node quadratic triangular element. The nodal

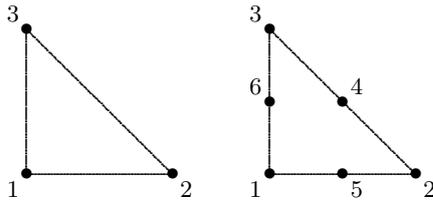


Figure B.4: Linear and quadratic triangular elements.

point coordinates of the iso-parametric linear element are  $(\xi, \eta)_1 = (0.0, 0.0)$ ,  $(\xi, \eta)_2 = (1.0, 0.0)$  and  $(\xi, \eta)_3 = (0.0, 1.0)$ . Associated to these nodes the linear shape functions read

$$N_1 = 1 - \xi - \eta, \quad N_2 = \xi, \quad N_3 = \eta. \quad (\text{B.33})$$

The coordinates of the nodes added by the quadratic element are  $(\xi, \eta)_4 = (0.5, 0.5)$ ,  $(\xi, \eta)_5 = (0.5, 0.0)$  and  $(\xi, \eta)_6 = (0.0, 0.5)$ , and the quadratic shape functions read

$$\begin{aligned} N_1 &= (1 - 2\xi - 2\eta)(1 - \xi - \eta), & N_2 &= \xi(2\xi - 1), & N_3 &= \eta(2\eta - 1), \\ N_4 &= 4\xi\eta, & N_5 &= 4\xi(1 - \xi - \eta), & N_6 &= 4\eta(1 - \xi - \eta). \end{aligned} \quad (\text{B.34})$$

Finally figure B.5 presents the linear and quadratic tetrahedral elements. The

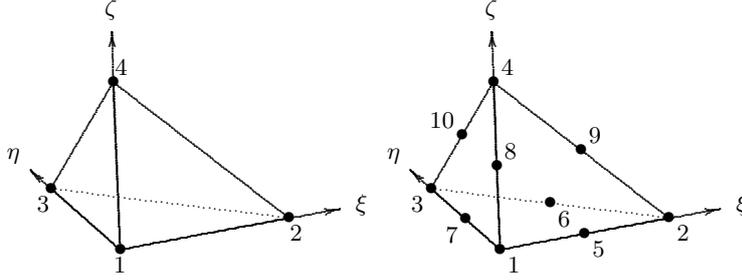


Figure B.5: Linear and quadratic tetrahedral element.

nodal point coordinates of the linear tetrahedral element are given by  $(\xi, \eta, \zeta)_1 = (0.0, 0.0, 0.0)$ ,  $(\xi, \eta, \zeta)_2 = (1.0, 0.0, 0.0)$ ,  $(\xi, \eta, \zeta)_3 = (0.0, 1.0, 0.0)$  and  $(\xi, \eta, \zeta)_4 = (0.0, 0.0, 1.0)$ . The shape functions then read

$$N_1 = 1 - \xi - \eta - \zeta, \quad N_2 = \xi, \quad N_3 = \eta, \quad N_4 = \zeta. \quad (\text{B.35})$$

The coordinates of the extra nodes introduced by the quadratic element are  $(\xi, \eta, \zeta)_5 = (0.5, 0.0, 0.0)$ ,  $(\xi, \eta, \zeta)_6 = (0.5, 0.5, 0.0)$ ,  $(\xi, \eta, \zeta)_7 = (0.0, 0.5, 0.0)$ ,  $(\xi, \eta, \zeta)_8 = (0.0, 0.0, 0.5)$ ,  $(\xi, \eta, \zeta)_9 = (0.5, 0.0, 0.5)$  and  $(\xi, \eta, \zeta)_{10} = (0.0, 0.5, 0.5)$ . The shape functions for the quadratic tetrahedral element are given by

$$\begin{aligned} N_1 &= (1 - 2\xi - 2\eta - 2\zeta)(1 - \xi - \eta - \zeta), & N_2 &= \xi(2\xi - 1), & N_3 &= \eta(2\eta - 1), \\ N_4 &= \zeta(2\zeta - 1), & N_5 &= 4\xi(1 - \xi - \eta - \zeta), & N_6 &= 4\xi\eta, & N_7 &= 4\eta(1 - \xi - \eta - \zeta), \\ N_8 &= 4\zeta(1 - \xi - \eta - \zeta), & N_9 &= 4\xi\zeta, & N_{10} &= 4\eta\zeta. \end{aligned} \quad (\text{B.36})$$

### Multilevel basis functions

The single level basis functions, which were presented in the previous subsection, extend to multilevel basis function. Figure B.6 shows a one-dimensional parent element and its two child elements. Node 3 which is present in the refined element case can be used in the definition of the modified basis function of the parent element. Figure B.7 depicts the element hierarchy. The coarse scale element denoted by its nodes (1, 2) is divided into two fine scale elements (3, 2) and (1, 3). The definition of the basis functions for the fine scale element in coarse scale coordinates outlines the modification of the basis functions. The basis functions for the coarse scale line element with local nodal coordinates  $(\xi)_1 = (0)$  and  $(\xi)_2 = (1)$  were already given by

$$N_1^{k-1} = 1 - \xi, \quad N_2^{k-1} = \xi. \quad (\text{B.37})$$



Figure B.6: Parent line element and child elements.

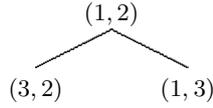


Figure B.7: Line element hierarchy.

Here  $k-1$  denotes the coarse scale. For the first fine scale element with nodal locations  $(\xi)_1 = (0.0)$  and  $(\xi)_3 = (0.5)$ , the basis functions read

$$N_1^k = 1 - 2\xi, \quad N_3^k = 2\xi. \quad (\text{B.38})$$

For the second fine scale element with nodal locations  $(\xi)_3 = (0.5)$  and  $(\xi)_2 = (1.0)$ , the basis functions read

$$N_2^k = 2\xi - 1, \quad N_3^k = 2 - 2\xi. \quad (\text{B.39})$$

Standard coarse scale basis functions  $N_a^{k-1}$  can now be written as the sum of fine scale basis functions  $N_a^k$ . For the linear line element this expression reads

$$N_1^{k-1} = N_1^k + \frac{1}{2}N_3^k, \quad N_2^{k-1} = N_2^k + \frac{1}{2}N_3^k. \quad (\text{B.40})$$

In matrix-vector notation this expression is given by  $[N_1^{k-1}] = [P]^T [N_1^k]$ . The interpolation matrix  $[P]$  reads

$$[P]^T = \begin{bmatrix} 1.0 & 0.0 & 0.5 \\ 0.0 & 1.0 & 0.5 \end{bmatrix}. \quad (\text{B.41})$$

The interpolation matrix displays the fine scale basis function weights. The multiscale finite element method aims at a modification of these weight such that it obtains fine scale accuracy on a coarse scale discretization. The multilevel formulation of basis functions also applies to higher-order elements and multidimensional elements.

The left picture of figure B.8 presents one triangular parent elements. The rightmost picture presents four child elements which are derived from an intermediate refinement phase for which two elements were generated. The hierarchy of elements is also presented by figure B.9. The coarse scale element  $(1, 2, 3)$  denoted by its nodes, is first divided into element  $(1, 4, 3)$  and  $(1, 2, 4)$ . The first of them generates  $(6, 4, 3)$  and  $(1, 4, 6)$  and the second intermediate element generates  $(5, 2, 4)$  and  $(1, 5, 4)$ . Standard coarse scale linear triangular elements basis functions follow from their fine scale counterparts as

$$\begin{aligned} N_1^{k-1} &= N_1^k + \frac{1}{2}N_5^k + \frac{1}{2}N_6^k, \\ N_2^{k-1} &= N_2^k + \frac{1}{2}N_4^k + \frac{1}{2}N_5^k, \\ N_3^{k-1} &= N_3^k + \frac{1}{2}N_4^k + \frac{1}{2}N_6^k. \end{aligned} \quad (\text{B.42})$$

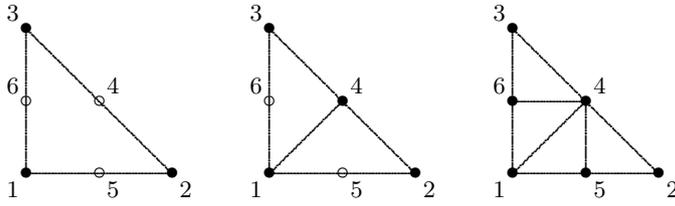


Figure B.8: Triangular parent and two-step child elements.

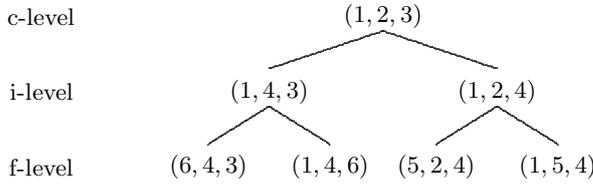


Figure B.9: Triangular element hierarchy.

Alternatively the coarse scale basis functions can be found by weighting the fine scale basis functions with the coefficients of the interpolation matrix. This matrix is given by

$$[P]^T = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.5 \\ 0.0 & 1.0 & 0.0 & 0.5 & 0.5 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.5 & 0.0 & 0.5 \end{bmatrix}. \tag{B.43}$$

Figure B.10 shows the refinement process of a tetrahedral parent element. In the

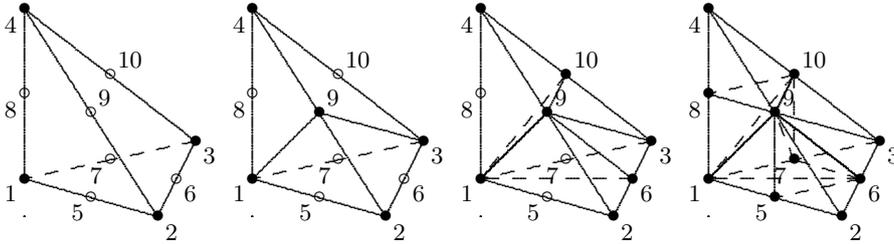


Figure B.10: Tetrahedral parent and child elements.

first step the coarse element  $(1, 2, 3, 4)$  generates two intermediate elements  $(1, 9, 3, 4)$  and  $(1, 2, 3, 9)$ . The second step splits these elements into  $(1, 9, 10, 4)$ ,  $(1, 9, 3, 10)$ ,  $(5, 2, 3, 9)$ , and  $(1, 5, 3, 9)$ . Finally the third step constructs the elements on the fine scale:  $(8, 9, 10, 4)$ ,  $(1, 9, 10, 8)$ ,  $(7, 9, 3, 10)$ ,  $(1, 9, 7, 10)$ ,  $(5, 6, 3, 9)$ ,  $(5, 2, 6, 9)$ ,  $(7, 5, 3, 9)$  and  $(1, 5, 7, 9)$ . Figure B.11 illustrates this process graphically. Standard coarse scale

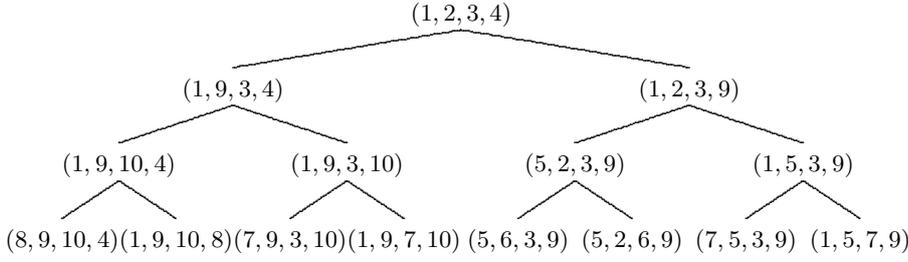


Figure B.11: Tetrahedral element hierarchy.

linear tetrahedral elements basis functions follow from their fine scale counterparts as

$$\begin{aligned}
 N_1^{k-1} &= N_1^k + \frac{1}{2}N_5^k + \frac{1}{2}N_7^k + \frac{1}{2}N_8^k, \\
 N_2^{k-1} &= N_2^k + \frac{1}{2}N_5^k + \frac{1}{2}N_6^k + \frac{1}{2}N_9^k, \\
 N_3^{k-1} &= N_3^k + \frac{1}{2}N_6^k + \frac{1}{2}N_7^k + \frac{1}{2}N_{10}^k, \\
 N_4^{k-1} &= N_4^k + \frac{1}{2}N_8^k + \frac{1}{2}N_9^k + \frac{1}{2}N_{10}^k.
 \end{aligned} \tag{B.44}$$

The interpolation matrix which relates the fine scale elements to the coarse scale parent reads

$$[P]^T = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 & 0.5 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 & 0.5 \\ 0.0 & 0.0 & 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.5 & 0.5 \end{bmatrix}. \tag{B.45}$$

### Integration in space

Integral functions need to be evaluated in the finite element formulation of partial differential equations. Integration of a spatial function  $f(\mathbf{x})$  over the two-dimensional element area  $\Omega^e$  can be written as

$$\int_{\Omega^e} f(\mathbf{x})d\Omega = \int_x \int_y f(x, y)dxdy = \int_{\xi} \int_{\eta} f(x(\xi, \eta), y(\xi, \eta))|J|d\xi d\eta, \tag{B.46}$$

where the Jacobian determinant supports a coordinate transformation. Numerical integration approximates this integral. Throughout this thesis numerical integration is carried out by Gaussian quadrature. In two directions the numerical approximation to the previous equation reads

$$\begin{aligned}
 &\int_{\xi} \int_{\eta} f(x(\xi, \eta), y(\xi, \eta))|J|d\xi d\eta \\
 &\approx \sum_{i=1}^{n_{m1}} \sum_{j=1}^{n_{m2}} w_i w_j f(x(\xi^{ij}, \eta^{ij}), y(\xi^{ij}, \eta^{ij}))|J|^{ij} = w_m (f|J|)_m, \tag{B.47}
 \end{aligned}$$

where  $n_{m1}$  is the number of integration points in  $\xi$  direction and  $n_{m2}$  is the number of integration points in  $\eta$  direction. In the short form  $n_m = n_{m1} \times n_{m2}$  is the total number of integration points, and  $w_m$  is the weight of integration point  $m$ . Both the function and the Jacobian determinant are evaluated at the local point  $m$ . For simplex elements, the Jacobian matrix is constant over the element.

Table B.1 presents the interpolation points, their position and weights for a general line element. According to this table, a single integration point calculates the integral of a linear function over the line element exactly. Two integration points obtain cubic accuracy. The sum of the weights equals one for all integration schemes, as the sum

Table B.1: BAR2 element, quadrature point information.

point	$\xi$	$w$	order
1	0.50000000	1.00000000	linear
1	0.21113249	0.50000000	cubic
2	0.78886751	0.50000000	

corresponds to the length of the isoparametric element.

Table B.2 assembles the integration points for the triangular element and Table B.3 gathers the locations of the quadrature points in a tetrahedral element and their weights. The integration points weights for triangular elements sum to 0.5 which

Table B.2: TRI3 element, quadrature point information.

point	$\xi$	$\eta$	$w$	order
1	0.33333333	0.33333333	0.50000000	linear
1	0.33333333	0.33333333	-0.28125000	cubic
2	0.20000000	0.20000000	0.26041667	
3	0.60000000	0.20000000	0.26041667	
4	0.20000000	0.60000000	0.26041667	

corresponds to the area of the local element. For tetrahedral elements the sum equals 0.167, this corresponds to the volume of the local element.

For linear accuracy the integration error reads  $R = O(h^2)$ , for cubic accuracy the error reads  $R = O(h^4)$ . Here  $h$  is the corresponding length of the mapped global element.

Table B.3: TET4 element, quadrature point information.

point	$\xi$	$\eta$	$\zeta$	$w$	order
1	0.25000000	0.25000000	0.25000000	0.16666667	linear
1	0.25000000	0.25000000	0.25000000	-0.13333333	cubic
2	0.16666667	0.16666667	0.16666667	0.07500000	
3	0.50000000	0.16666667	0.16666667	0.07500000	
4	0.16666667	0.50000000	0.16666667	0.07500000	
5	0.16666667	0.16666667	0.50000000	0.07500000	

### Integration in time

The finite element discretization in space generally results in a (non-linear) set of ordinary differential equations, which can be expressed as

$$M_{ab} \frac{du_b}{dt} = S_{ab}u_b + F_a, \quad (\text{B.48})$$

where  $u_b$  denote the components of the unknown,  $M_{ab}$  are the components of the capacity matrix,  $S_{ab}$  denotes the conductivity matrix components and  $F_a$  are the force vector components. Finite difference time integration schemes translate the set of (nonlinear) ordinary differential equations into a set of (nonlinear) algebraic equations. The result of a linear time integration scheme, can be written as

$$M_{ab}^{n+\theta} \frac{u_b^{n+1} - u_b^n}{\Delta t} = \theta S_{ab}^{n+1} u_b^{n+1} + (1 - \theta) S_{ab}^n u_b^n + \theta F_a^{n+1} + (1 - \theta) F_a^n. \quad (\text{B.49})$$

In this expression  $n + 1$  denotes the next time level,  $n$  is the current time level and  $\theta$  is a time-weighting factor for which  $0 \leq \theta \leq 1$ . The capacity matrix  $M_{ab}^{n+\theta}$  is obtained by a weighted average of its value in the current and the next time step. The conductivity matrix and force vector are evaluated separately at the current and next time step. Explicit time integration follows for  $\theta = 0$ , fully implicit time integration for  $\theta = 1$ , and second order accuracy time integration is obtained for  $\theta = \frac{1}{2}$ .

Large sets of equations are often integrated explicitly and the matrix is diagonalized by a process called mass lumping. The lumped capacity matrix approximates the consistent mass matrix as

$$\int_{\Omega_e} C N_a N_b d\Omega \approx \int_{\Omega_e} \delta_{ab} C N_a d\Omega. \quad (\text{B.50})$$

This procedure results in an explicit expression of the unknown approximation even for nonlinear ordinary differential equations.

If the capacity matrix depends on the unknown in time, and an implicit scheme was chosen ( $\theta > 0$ ), then the application of the chain rule gives

$$\frac{dM_{ab}}{du} \frac{du_b}{dt} = S_{ab}u_b + F_a. \quad (\text{B.51})$$

The partial derivative of the capacity matrix coefficients are obtained either tangentially to the point in time  $n + \theta$  or follow from their values at the current and new time step. The tangential form is written as

$$\frac{\partial M_{ab}}{\partial u_b} \Big|^{n+\theta} \frac{u_b^1 - u_b^0}{\Delta t} = \theta S_{ab}^{n+1} u_b^{n+1} + (1 - \theta) S_{ab}^n u_b^n + \theta F_a^{n+1} + (1 - \theta) F_a^n. \quad (\text{B.52})$$

The secant form is expressed as

$$\frac{M_{ab}(u_b^{n+1}) - M_{ab}(u_b^n)}{u_b^{n+1} - u_b^n} \frac{du_b}{dt} = \theta S_{ab}^{n+1} u_b^{n+1} + (1 - \theta) S_{ab}^n u_b^n + \theta F_a^{n+1} + (1 - \theta) F_a^n. \quad (\text{B.53})$$

In addition to numerical analysis, numerical experiments can show which form is preferred.

### Multiscale basis functions

Multiscale basis functions extend multilevel basis functions. Figures B.12, B.13, B.14 and B.15 present multiscale basis functions for a number of periodic fields. The functions capture the fine scale behavior of horizontally layered fields, checkerboard fields, domains with inclusions and diagonally layered fields on a patch of elements. Multiscale basis functions were obtained by two techniques. The first approach computes multiscale basis functions on an extended local domain. The domain is closed by linear boundary conditions, and includes a subdomain of elements that captures a single periodic cell. The second approach closes the local flow problems by oscillating boundary conditions, which follow from dimensionally reduced flow problems. The permeability contrast for all periodic problems equals  $10^2$  and the finite element mesh over-resolves the structure. For each periodic case the upper plots give the basis functions on the oversampled domain. The plots below restrict these basis functions to a subdomain according to the oversampling approach. The lower plots give the multiscale basis functions that follow from closure of the local flow problems by oscillating boundary conditions.

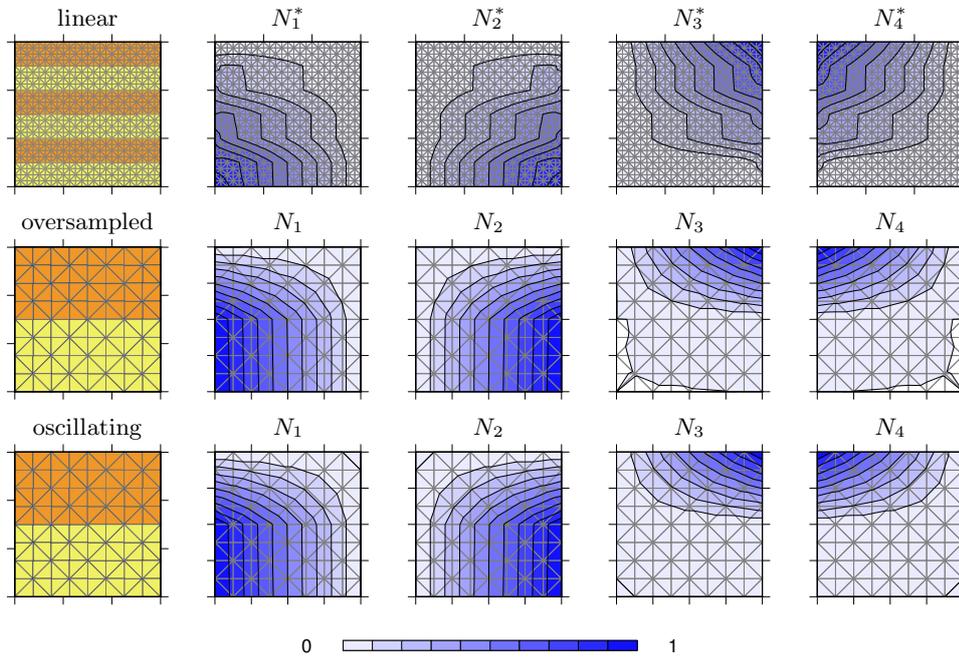


Figure B.12: Multiscale basis functions on a horizontally layered field.

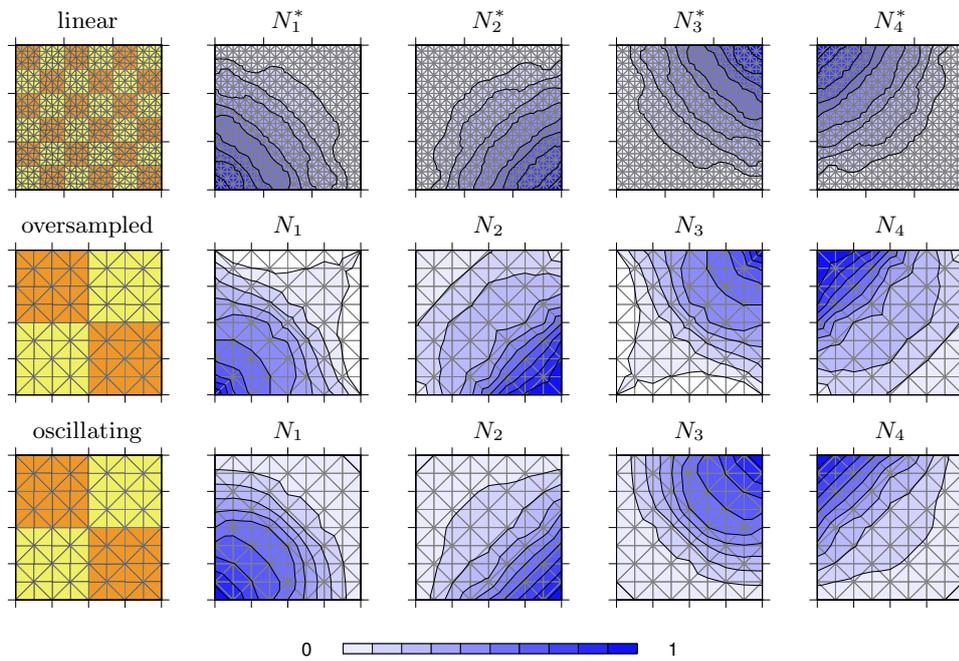


Figure B.13: Multiscale basis functions on a checkerboard field.

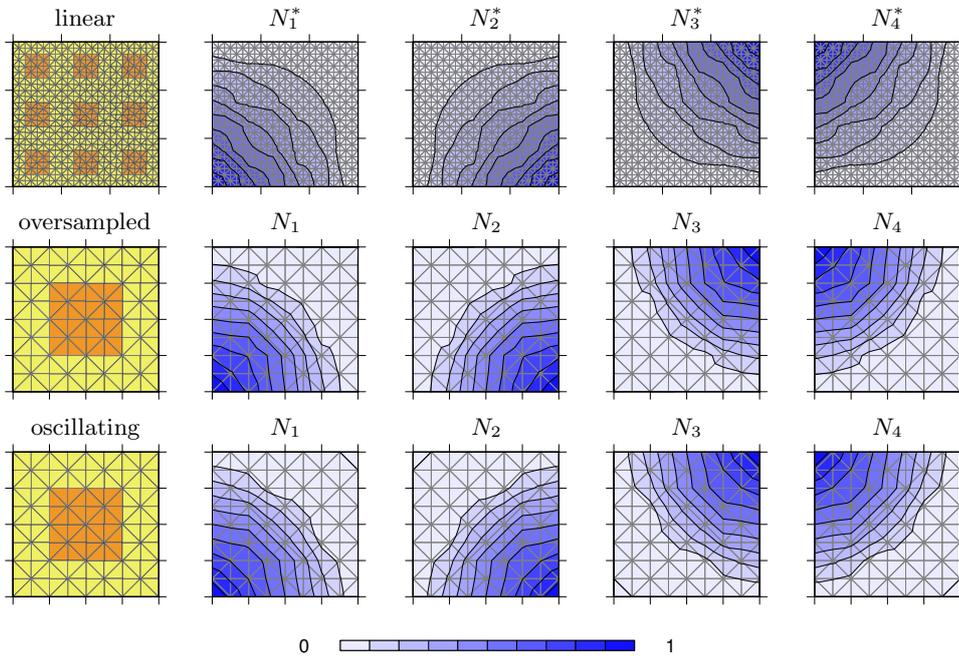


Figure B.14: Multiscale basis functions on a field with inclusions.

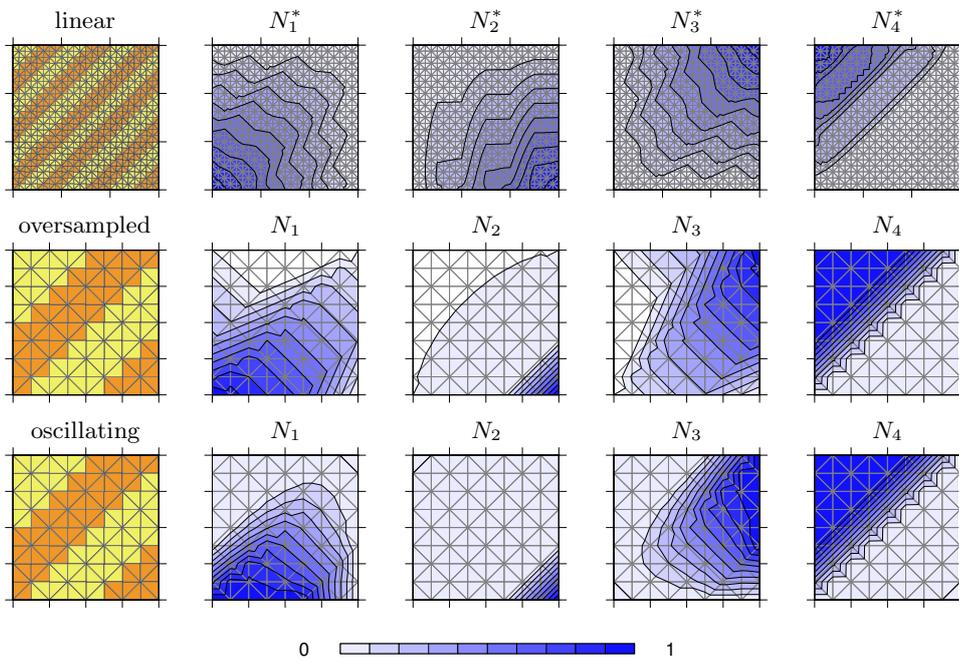


Figure B.15: Multiscale basis functions on a diagonally layered field.

# Appendix C

## Multigrid method

Numerical solutions for sets of algebraic equations are either obtained by direct numerical solvers or by iterative numerical solvers [104, 57]. Gauss-Jordan elimination provides a classical direct solution procedure. Basic iterative solvers are Jacobi's method and the Gauss-Seidel method. Conjugate gradient solvers and multigrid solvers are classified as more advanced iterative solvers, which aim to solve a set of  $n_u$  equations in  $O(n_u)$  operations [60, 110, 122, 18]. Multigrid solvers improve basic iterative solvers, and accelerate the relaxation of smooth components of the fine-grid residual by solving the residuals in the unknowns at a coarser grid.

This appendix introduces a linear multigrid algorithm by a two-grid correction scheme. The nonlinear multigrid algorithm then generalizes the linear concept. Special techniques must be applied to restore the efficiency of the multigrid solver for problems with highly heterogeneous field parameters. A number of algebraic multigrid techniques will be given, which construct an (approximate) Schur complement with respect to the coarse grid. The Schur complement provides a separation of scales, as coarse-grid unknowns are related to other coarse-grid unknowns only. Gauss-Jordan elimination introduces the Schur complement, which provides the exact coarse-scale operator and optimal interpolation operator. However, a global system of equations needs to be inverted to obtain the Schur complement in multi-dimensional cases. The construction of an approximated matrix overcomes this problem. For quadrilateral elements a directional lumping procedure is often applied.

### Multigrid algorithms

Basic iterative solvers like Gauss-Seidel shows poor convergence behavior as the number of iterations increase. The Gauss-Seidel iteration scheme for a lexicographic ordered mesh, presented by the picture left in figure C.1, can be written as

$$u_i^{r+1} = -\frac{1}{L_{ii}} \sum_{j=1}^{i-1} L_{ij} u_j^{r+1} - \frac{1}{L_{ii}} \sum_{j=i+1}^{n_u} L_{ij} u_j^r + \frac{1}{L_{ii}} f_i, \quad (\text{C.1})$$

where  $n_u$  is the number of unknowns, and summation only takes place over  $j$ , and  $i = 1, \dots, n_u$ . The algorithm takes updated values of the unknowns into account,

as soon as they become available. In matrix-vector form this equation reads  $[u]^1 = [R_G][u]^0 + [L^D]^{-1} + [f]$ , where  $[R_G]$  is the iteration matrix and  $[L^D]$  denotes the diagonal matrix. For the Gauss-Seidel method the iteration matrix reads

$$[R_G] = ([L^D] - [L^U])^{-1} [L^L]. \quad (\text{C.2})$$

Here  $[L^U]$  is the upper matrix and  $[L^L]$  is the lower matrix. The performance of the

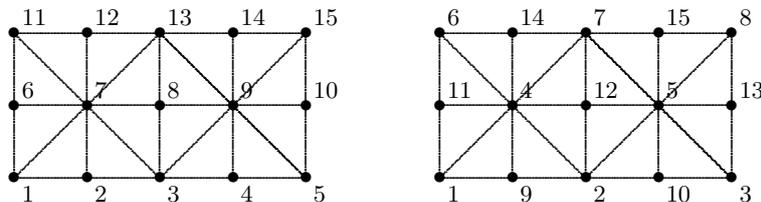


Figure C.1: Node ordering schemes.

Gauss-Seidel method depends on the mesh ordering. A red-black ordering of the mesh points improves the convergence speed of the Gauss-Seidel method. The picture right in figure C.1 presents this alternative node ordering. Jacobi's method follows from the Gauss-Seidel method if no updates are used for the unknowns, by this Jacobi's method becomes mesh independent. Both Jacobi's method and Gauss-Seidel method take into account only the direct neighboring nodes of the unknown  $u_i$ , and therefore they are called local operations.

The effectiveness of the scheme in reducing the error depends on the spatial frequency of the error components. For basic iterative methods, the convergence behavior deteriorates as iterations continue. The reason for this becomes clear if a discrete Fourier analysis of the error is elaborated. The relaxation process is very effective on error components whose periods are small, but is ineffective in smoothing the low-frequency components of the error. For solving a discrete boundary value problem on a mesh with characteristic size  $h$ , periods of a few  $h$  are called high-frequency errors, and errors with periods greater than about  $10h$  are called low frequency errors. Multigrid methods use a sequence of coarser meshes to speed up their convergence. In a few iterations basic relaxation methods reduce the high frequency errors on the fine mesh. The remaining error is then transferred to a coarser grid where it can be reduced more effectively, as low frequency error components on a fine mesh become high frequency components on a coarse mesh.

### Linear multigrid algorithm

The basic idea of the linear multigrid algorithm will be introduced on two grid levels, with uniform grid sizes  $h$  and  $H$ , and standard coarsening  $H = 2h$ . The continuous unknown will be written as  $u$  and its discrete approximation on the grids is denoted as  $u_h$  and  $u_H$ . As the multigrid method is an iterative technique, the sequential solution at iteration  $r$  will be written as  $u_h^r$  and  $u_H^r$ . An iteration from the fine level to the coarsest level and back to the fine level, called cycle, produces the next iterate.

The continuous boundary value problem that resembles a linear flow problem will be written symbolically as

$$\mathcal{L}(u) = f \quad \text{on } \Omega, \quad (\text{C.3})$$

where  $\mathcal{L}$  is the continuous linear elliptic differential operator and  $f$  is the source function. On a given uniform grid  $\Omega_h$  with mesh size  $h$ , its discrete form is written as

$$\mathcal{L}_h(u_h) = f_h \quad \text{on } \Omega_h, \quad (\text{C.4})$$

with linear elliptic operator matrix  $\mathcal{L}_h$ , source vector  $f_h$ , and discrete approximation to the continuous equation  $u_h$ . The jet unknown exact solution  $u_h$  to the set of algebraic equations will be approximated by  $u_h^r$ .

The error in the approximate solution or correction  $e_h^r$  reads

$$e_h^r = u_h - u_h^r. \quad (\text{C.5})$$

The defect correction scheme improves the approximate solution  $u_h^r$  at level  $h$  in iteration  $r$ . The defect or residual  $d_h^r$  follows from

$$d_h^r = f_h - \mathcal{L}_h(u_h^r). \quad (\text{C.6})$$

Since  $\mathcal{L}_h$  is linear,  $\mathcal{L}_h(u_h^r) = \mathcal{L}_h(u_h) - \mathcal{L}_h(e_h^r)$  holds and therefore the error satisfies

$$\mathcal{L}_h(e_h^r) = d_h^r. \quad (\text{C.7})$$

This equation is known as the defect equation. The approximate solution for the error, written as  $\hat{e}_h^r$ , can be found from

$$\hat{\mathcal{L}}_h(\hat{e}_h^r) = d_h^r, \quad (\text{C.8})$$

where  $\hat{\mathcal{L}}_h$  is an approximated operator. Multigrid techniques use the approximation of the operator on the next coarse level  $\mathcal{L}_H$  to calculate the error at that level as

$$\mathcal{L}_H(e_H^r) = d_H^r. \quad (\text{C.9})$$

This coarse-scale equation has a smaller dimension and is easier to solve than the fine-scale problem. To obtain a solution for the error  $e_H^r$  on the coarse-scale, first the defect of the fine grid has to be transferred to the coarse grid. The defect  $d_H^r$  on the coarse grid is found from an intergrid transfer operation which is written as

$$d_H^r = \mathcal{R}(d_h^r), \quad (\text{C.10})$$

here  $\mathcal{R}$  denotes the restriction operator. This operator is also known as fine-to-coarse operator, injection operator or extrapolation operator. After solving the defect equation at the coarse grid, the solution of  $e_H^r$  has to be transferred back to the fine grid. The corresponding error  $\hat{e}_h^r$  at the fine scale is found by interpolation as

$$\hat{e}_h^r = \mathcal{P}(e_H^r), \quad (\text{C.11})$$

where  $\mathcal{P}$  denotes the prolongation operator, which is also known as coarse-to-fine operator or interpolation operator. Finally the corrected value of the unknown at the fine mesh nodes, is found as

$$u_h^{r+1} = u_h^r + \hat{e}_h^r. \quad (\text{C.12})$$

A two-level multigrid iteration, or cycle, results from the two-grid correction step if it is preceded by a pre-smoothing step on the fine level and completed by a post-smoothing operation on the fine grid. Figure C.2 presents a two grid cycle graphically.

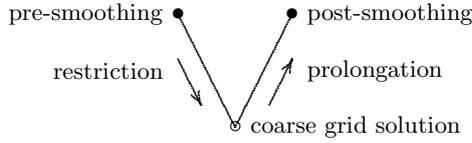


Figure C.2: Two-grid cycle.

Equation C.13 specifies the call of a single cycle of the linear two-grid method  $LTG(\cdot)$ , which produces a new fine-scale approximation of the unknowns  $[u_h^{r+1}]$  given an initial guess of the unknowns  $[v_h] = [u_h^r]$ . Here  $\nu_1$  accounts for the number of pre-smoothing operations and  $\nu_2$  denotes the number of post-smoothing operations.

$$[u_h] = LTG([v_h], [L_h], [f_h], \nu_1, \nu_2) \quad (\text{C.13})$$

- relax  $\nu_1$  times on  $[L_h][u_h] = [f_h]$  with initial guess  $[u_h] = [v_h]$
- compute the fine-grid defect  $[d_h] = [f_h] - [L_h][u_h]$
- restrict the fine-scale defect to the coarse grid  $[d_H] = [R_h^H][d_h]$
- solve the defect equation  $[L_H][e_H] = [d_H]$
- prolongate the coarse-grid correction to the fine grid  $[\hat{e}_h] = [P_H^h][e_H]$
- correct the fine-scale approximation  $[v_h] = [u_h] + [\hat{e}_h]$
- relax  $\nu_2$  times on  $[L_h][u_h] = [f_h]$  with initial guess  $[u_h] = [v_h]$

Given an initial guess  $[u_h^0]$  the first cycle produces  $[u_h^1] = LTG([u_h^0], \dots)$ , the second cycle gives  $[u_h^2] = LTG([u_h^1], \dots)$  and so on. The two-grid method is easily extended to a three-grid method if the defect equation is solved by a recursive call of the two-grid method. General multigrid methods solve a nested sequence of algebraic equations, which is labeled by  $k = 1, \dots, k_f$ . Here  $k = 1$  indicates the coarsest grid and  $k_f$  indicates the finest grid. Equation (C.14) generalizes the two-grid method, and gives a recursive formulation of the linear multigrid algorithm  $LMG(\cdot)$  for a single cycle. Relaxing  $\nu_1$  times on  $[L^k][u^k] = [f^k]$  with initial guess  $[u^k] = [v^k]$  has now been replaced by the expression  $[u^k] = LS([v^k], [L^k], [f^k], \nu_1, \nu_2)$ .

On three grid levels, the call  $[u^3] = LMG(3, \dots)$  on the finest grid for  $k = 3$  carries out a pre-smoothing operation on  $[u^3]$ . It then proceeds to a coarser scale in the downward stroke, and calls the linear multigrid operator  $[u^2] = LMG(2, \dots)$  before

it proceeds to a finer scale along the upward stroke. This recursive call sets  $k = 2$  and performs a pre-smoothing operation on  $[u^2]$ , then solves the linear equation at the coarsest grid  $[u^1] = LS(1, \dots)$ , and performs a post-smoothing operation on  $[u^2]$ . The algorithm leaves the recursive call and proceeds with a post-smoothing operation on  $[u^3]$ .

$$[u^k] = LMG(k, [v^k], [L^k], [f^k], \nu_1, \nu_2) \quad (\text{C.14})$$

- downward stroke
  - pre-smoothing  $[u^k] = LS([v^k], [L^k], [f^k], \nu_1)$
  - compute defect  $[d^k] = [f^k] - [L^k][u^k]$
  - restrict defect  $[d^{k-1}] = [R_k^{k-1}][d^k]$
  - introduce  $[v^{k-1}] = [0]$
- solve coarse grid equation
  - if  $k > 2$ :  $[u^{k-1}] = LMG(k-1, [v^{k-1}], [L^{k-1}], [d^{k-1}], \nu_1, \nu_2)$
  - if  $k = 2$ :  $[u^{k-1}] = LS([v^{k-1}], [L^{k-1}], [d^{k-1}], \nu_1, \nu_2)$
- upward stroke
  - prolongate correction  $[\hat{e}^k] = [P_{k-1}^k][u^{k-1}]$
  - compute new approximation  $[w^k] = [u^k] + [\hat{e}^k]$
  - post-smoothing  $[u^k] = LS([w^k], [L^k], [f^k], \nu_2)$

The discrete differential operator at the finest level follows from a finite element discretization of the flow equation. Its matrix representation is given by  $[L^{k_f}]$ . The right hand side vector is not included in the operator and written separately as  $[f^{k_f}]$ . Equation (C.14) provides coarse-grid right-hand side vectors. At the coarser levels  $k = 1, \dots, k_{f-1}$  however, coarse-grid operators have to be calculated and intergrid transfer operators have to be quantified. The matrix representation of the discrete prolongation operator is denoted by  $[P_{k-1}^k]$  as it interpolates the result from grid  $k-1$  to grid  $k$ . The matrix representation of the restriction operator is given by  $[R_k^{k-1}]$  as it transfers the results from level  $k$  to  $k-1$ . The extra computational effort of the multigrid method over basic iterative methods is limited. For a  $n$ -dimensional problem the relaxation work and memory requirement decrease by a factor  $1/2^n$  each time the grid spacing is doubled. Accordingly, for a three-dimensional problem the memory requirement only increases with about 15% as  $1 + 1/8 + 1/64 + \dots \approx 1.143$ .

### Non-linear multigrid algorithm

The nonlinear multigrid algorithm known as the Full Approximation Storage (FAS) algorithm [99] extends the linear multigrid algorithm. Press et al. [99] presented two complementary viewpoints for the relation between coarse and fine grids for the FAS algorithm. The first viewpoint poses that coarse grids are used to accelerate

the convergence of the smooth components of the fine-grid residuals. The second viewpoint states that fine grids are used to compute correction terms to the coarse-grid equations. The correction terms project fine-grid accuracy to the coarse grid. The FAS algorithm provides a stopping criterion and an approximate solution at each grid level. If the algorithm applies nonlinear smoothing operators, it is capable to solve nonlinear equations. The nonlinear algorithm generalizes the linear algorithm.

The nonlinear continuous boundary value problem is written symbolically as

$$\mathcal{N}(u) = f \quad \text{on } \Omega_h. \quad (\text{C.15})$$

The discrete boundary value problem on a uniform grid  $h$  reads

$$\mathcal{N}_h(u_h) = f_h \quad \text{on } \Omega_h. \quad (\text{C.16})$$

The error in the discrete equation will be written as

$$e_h = u - u_h. \quad (\text{C.17})$$

The two-grid correction scheme uses the approximate solution  $u_h^r$  of iteration  $r$  at the fine level  $h$ . Following the first viewpoint, the error in the approximate solution  $e_h^r$  reads

$$e_h^r = u_h - u_h^r, \quad (\text{C.18})$$

where  $u_h$  is the unknown exact discrete solution at mesh  $h$ . The discrete equation can now be written as

$$\mathcal{N}_h(u_h^r + e_h^r) = f_h. \quad (\text{C.19})$$

The defect equation follows from

$$\mathcal{N}_h(u_h^r + e_h^r) - \mathcal{N}_h(u_h^r) = d_h^r, \quad (\text{C.20})$$

where  $d_h^r = f_h - \mathcal{N}_h(u_h^r)$ . The approximate solution on a coarse level is found by solving

$$\mathcal{N}_H(u_H^r) = \mathcal{N}_H(\mathcal{R}(u_h^r)) + \mathcal{R}(d_h^r). \quad (\text{C.21})$$

Now the coarse-grid correction reads

$$e_H^r = u_H^r + \mathcal{R}(u_h^r). \quad (\text{C.22})$$

The coarse-grid correction modifies the solution at the fine scale according to

$$u_h^{r+1} = u_h^r + \mathcal{P}(e_H^r). \quad (\text{C.23})$$

This process is repeated to obtain the next iterate  $u_h^{r+2}$ , and so on. A stopping criterion could be based on the defect norm, written as

$$\|d_h\| \leq \epsilon. \quad (\text{C.24})$$

A reasonable choice of  $\epsilon$  can be found from the second viewpoint.

This dual view point considers local truncation errors  $\tau_h$  that can be regarded as corrections to the discrete source terms  $f_h$

$$\mathcal{N}_h(u) = f_h + \tau_h, \quad (\text{C.25})$$

where  $u$  is the exact solution of the continuum equation. The relative truncation error that corrects the discrete equation at grid size  $h$  to the continuous equation now reads  $\tau_h = \mathcal{N}_h(u) - f_h$ . The truncation error defined on the coarse-grid relative to the fine grid is written as

$$\tau_H = \mathcal{N}_H(\mathcal{R}(u_h)) - \mathcal{R}(\mathcal{N}_h(u_h)). \quad (\text{C.26})$$

Here  $u_h$  is the unknown exact discrete solution on the fine scale. The relative truncation error  $\tau_H$  corrects the coarse-scale source term  $f_H$  and makes the solution of the coarse-grid equation equal to the solution of the fine-grid equation. The corrected coarse-scale equations are expressed by

$$\mathcal{N}_H(u_H) = f_H + \tau_H. \quad (\text{C.27})$$

An approximation of  $\tau_H$  follows from the approximate solution on the fine scale  $u_h^r$  and reads

$$\tau_H^r = \mathcal{N}_H(\mathcal{R}(u_h^r)) - \mathcal{R}(\mathcal{N}_h(u_h^r)). \quad (\text{C.28})$$

Substitution of equation (C.28) into equation (C.27) gives

$$\mathcal{N}_H(u_H^r) = \mathcal{N}_H(\mathcal{R}(u_h^r)) + [f_H - \mathcal{R}(\mathcal{N}_h(u_h^r))], \quad (\text{C.29})$$

where the last two terms form the defect  $d_H$ . If the coarse-grid force vector is obtained by reduction as  $f_H = \mathcal{R}(f_h)$ , then equation (C.29) gives equation (C.21). For a typical case of second-order accurate differencing Press et al. [99] found  $\tau \simeq 1/3\tau_h^r$  and hence a good criterion is

$$\epsilon = \frac{1}{3} \|\tau_h^r\|. \quad (\text{C.30})$$

This choice of  $\epsilon$  prevents iterating beyond the point when the remaining error is dominated by the local truncation error.

Equation (C.31) gathers the two-grid nonlinear multigrid scheme. The nonlinear two-grid algorithm  $NTG(\cdot)$  produces a new fine-scale approximation  $[u_h]$ , given an initial guess of the unknowns  $[v_h]$ . It carries out  $\nu_1$  pre-smoothing operations and  $\nu_2$  post-smoothing operations. The matrix representation of the nonlinear operator  $\mathcal{N}_h$  is written as  $[N_h]$ . The nonlinear two-grid algorithm  $NTG(\cdot)$  restricts the fine-grid approximation and the fine-grid defect. It solves the coarse-grid problem and obtains coarse-grid approximations. Finally it computes the fine-grid correction as the difference between the restricted fine-grid approximation and the calculated coarse-grid approximation. The linear two-grid algorithm  $LTG(\cdot)$  restricted the fine-grid defect only.  $LTG(\cdot)$  computed the coarse-grid correction by solving the defect equation. It did not calculate coarse-grid approximations; coarse-grid corrections were calculated

instead.

$$[u_h] = NTG([v_h], [N_h], [f_h], \nu_1, \nu_2) \quad (\text{C.31})$$

- relax  $\nu_1$  times on  $[N_h][u_h] = [f_h]$  with initial guess  $[u_h] = [v_h]$
- restrict the fine-grid approximation to the coarse grid  $[v_H] = [R_h^H][u_h]$
- compute the fine-grid defect  $[d_h] = [f_h] - [N_h][u_h]$
- restrict the fine-grid defect to the coarse grid  $[d_H] = [R_h^H][d_h]$
- solve the coarse-grid problem  $[N_H][u_H] = [N_H][v_H] + [d_H]$
- compute the coarse-grid correction  $[e_H] = [u_H] - [v_H]$
- prolongate the coarse-grid correction to the fine grid  $[\hat{e}_h] = [P_H^h][e_H]$
- correct the fine-scale approximation  $[v_h] = [u_h] + [\hat{e}_h]$
- relax  $\nu_2$  times on  $[N_h][u_h] = [f_h]$  with initial guess  $[u_h] = [v_h]$

The multigrid generalization is presented by equation (C.32). This equation presents the recursive formulation of the nonlinear multigrid iterator  $NMG(\cdot)$  for a single V-cycle, depicted by figure C.2, based on the FAS algorithm. The nested sequence of algebraic equations is labeled  $k = 1, \dots, k_f$  where 1 indicates the coarsest grid and  $k_f$  is the finest grid.

In general the nonlinear multigrid algorithm uses a nonlinear smoother  $NS(\cdot)$  to smooth high frequency errors on a fine mesh, whereas the linear multigrid algorithm used a linear smoother  $LS(\cdot)$ . The FAS algorithm presented by equation (C.32) needs one extra set of vectors to store the right hand side. This was not needed for the  $LMG(\cdot)$  algorithm given by equation (C.14). For a linear operator the nonlinear algorithm resembles the linear algorithm.

The formulation of the linear multigrid iterator  $LMG(\cdot)$  and the nonlinear multigrid iterator  $NMG(\cdot)$  was presented for a V-cycle schedule. Both methods can easily be extended to alternative schedules, which are computationally more effective occasionally. Figure C.3 presents the V-cycle next to the W-cycle and the F-cycle. All cycles should be carried out in sequence until a stopping criterion is met.

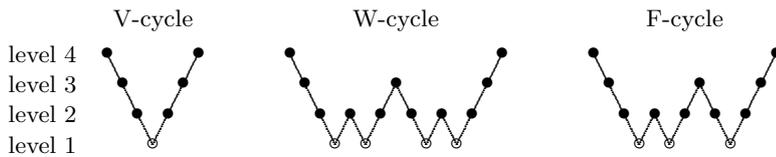


Figure C.3: Four level cycle schemes.

The efficiency of multigrid is improved by a procedure called 'nested iteration' for certain applications. The full multigrid algorithm or nested iteration procedure forms an extension of the linear multigrid algorithm. The full multigrid algorithm starts on the coarsest mesh, and needs a discrete force vector as well as a discrete coarse-grid

operator. A F-cycle scheme, without pre-smoothing operations on the first decent supplemented by zero initial solution, produces a full multigrid.

### Multigrid operators

The coarse-grid operator for the model problem with homogeneous field functions, simply follow from a finite element discretization on the coarse level. However, for multi-dimensional problems the intergrid transfer operations already pose a problem as fine scale unknowns relate to coarse scale unknowns and other fine scale unknowns. For this reason a concept of separation of variables needs to be posed.

$$[u^k] = NMG(k, [v^k], [N^k], [f^k], \nu_1, \nu_2) \quad (C.32)$$

- downward stroke
  - pre-smoothing  $[u^k] = NS([v^k], [N^k], [f^k], \nu_1)$
  - compute defect  $[d^k] = [f^k] - [N^k] [u^k]$
  - restrict defect  $[d^{k-1}] = [R_k^{k-1}] [d^k]$
  - restrict approximation  $[v^{k-1}] = [R_k^{k-1}] [u^k]$
  - compute right hand side  $[g^{k-1}] = [N^{k-1}] [v^{k-1}] + s^{k+1} [d^{k-1}]$
- solve coarse-grid equation
  - if  $k > 2$ :  $[u^{k-1}] = NMG(k-1, [v^{k-1}], [N^{k-1}], [g^{k-1}], \nu_1, \nu_2)$
  - if  $k = 2$ :  $[u^{k-1}] = NS([v^{k-1}], [N^{k-1}], [g^{k-1}], )$
- upward stroke
  - compute correction  $[e^{k-1}] = [u^{k-1}] - [v^{k-1}]$
  - prolongate correction  $[\hat{e}^k] = [P_{k-1}^k] [e^{k-1}]$
  - compute new approximation  $[w^k] = [u^k] + \frac{1}{s^{k-1}} [\hat{e}^k]$
  - post-smoothing  $[u^k] = NS([w^k], [N^k], [f^k], \nu_2)$

Gauss-Jordan elimination and the construction of the Schur complement introduce algebraic multigrid. The solution of a set of algebraic equations with two unknowns can easily be found by Gauss-Jordan elimination. The set of equations is written as

$$\begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}. \quad (C.33)$$

Subsequently multiplying the second row by  $m_1 = L_{12}/L_{22}$ , and subtracting the result from the first row, restates the set of equations as

$$\begin{bmatrix} L_{11} - L_{21}L_{12}/L_{22} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} f_1 - m_1 f_2 \\ f_2 \end{bmatrix}. \quad (C.34)$$

Multiplying the first row by  $m_2 = (L_{21}L_{22}) / (L_{11}L_{22} - L_{12}L_{21})$  and subtracting the result from the second row gives

$$\begin{bmatrix} L_{11} - L_{21}L_{12}/L_{22} & 0 \\ 0 & L_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} f_1 - m_1f_2 \\ f_2 - m_2(f_1 - m_1f_2) \end{bmatrix}. \quad (\text{C.35})$$

This direct solution introduces the concept of separation of variables, but it becomes inefficient for large sets of equations.

### Schur complement

Figure C.4 presents a coarse mesh with nodes belonging to  $\Omega^{k-1}$  and its refined mesh with nodes belonging to  $\Omega^k$ . In a multigrid setting, coarse meshes are available at grid levels  $k = 1, \dots, k_f - 1$ . At the fine mesh, a global coarse-fine node ordering (c-f ordering) is proposed. The set of equations which follows from a discretization at the fine mesh will now be rewritten as two separate sets: one set for the fine nodes also present in the coarse mesh (c-nodes) and one set for the fine-grid nodes that do not appear in the coarse grid (f-nodes). For the fine grid shown in figure C.4, nodal point

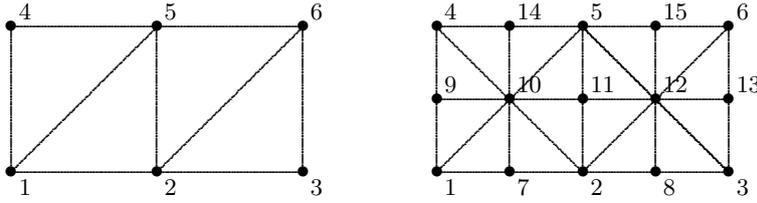


Figure C.4: Coarse mesh and c-f ordered fine mesh.

$1, \dots, 6$  are c-nodes belonging to  $\Omega^{k-1} \cap \Omega^k$ , and nodal points  $7, \dots, 15$  are f-nodes belonging to  $\Omega^k \setminus \Omega^{k-1}$ .

For a discrete problem  $L_{ab}^k u_b^k = f_a^k$  on a grid  $\Omega^k$ , the linear set of equations can be partitioned in coarse-grid entries belonging to  $\Omega^{k-1}$  and fine-grid entries that do not belong to the coarse-grid part  $\Omega^k \setminus \Omega^{k-1}$  as

$$\begin{bmatrix} [L^{11}] & [L^{12}] \\ [L^{21}] & [L^{22}] \end{bmatrix} \begin{bmatrix} [u^1] \\ [u^2] \end{bmatrix} = \begin{bmatrix} [f^1] \\ [f^2] \end{bmatrix}, \quad (\text{C.36})$$

where  $[u^1]$  collects the unknowns in c-points and  $[u^2]$  gathers the f-points unknowns. The block matrix  $[L^{11}]$  collects c-point to c-point connectivities and the block matrix  $[L^{22}]$  contains f-point to f-point connectivities. The optimal matrix-dependent prolongation arises from the exact block Gaussian elimination of the outer diagonal blocks  $[L^{12}]$  and  $[L^{21}]$  which collect f-node to c-node connectivities and vice versa. This

elimination process reads

$$\begin{aligned} \begin{bmatrix} [I] & -[L^{12}][L^{22}]^{-1} \\ [0] & [I] \end{bmatrix} \begin{bmatrix} [L^{11}] & [L^{12}] \\ [L^{21}] & [L^{22}] \end{bmatrix} \begin{bmatrix} [I] & [0] \\ -[L^{22}]^{-1}[L^{21}] & [I] \end{bmatrix} \\ = \begin{bmatrix} [L^{11}] - [L^{12}][L^{22}]^{-1}[L^{21}] & [0] \\ [0] & [L^{22}] \end{bmatrix}. \end{aligned} \quad (\text{C.37})$$

Equation (C.37) presents coarse-grid interactions by the matrix

$$[L^{k-1}] = [L^{11}] - [L^{12}][L^{22}]^{-1}[L^{21}]. \quad (\text{C.38})$$

This matrix forms the Schur complement of  $[L^k]$  with respect to  $\Omega^{k-1}$ . For a system of two unknowns the scalar form was found by Gauss-Jordan elimination  $L^{k-1} = L_{11}^k - L_{21}^k L_{12}^k / L_{22}^k$ .

Alternatively, the same result can be found by variational coarsening. The Galerkin condition poses the first variational property and reads

$$[L^{k-1}] = [R_k^{k-1}][L^k][P_{k-1}^k], \quad (\text{C.39})$$

with a restriction operator  $[R_k^{k-1}]$  given by

$$[R_k^{k-1}] = \begin{bmatrix} [I] & -[L^{12}][L^{22}]^{-1} \end{bmatrix}, \quad (\text{C.40})$$

and a prolongation operator  $[P_{k-1}^k]$  expressed as

$$[P_{k-1}^k] = \begin{bmatrix} [I] \\ -[L^{22}]^{-1}[L^{21}] \end{bmatrix}. \quad (\text{C.41})$$

For symmetric systems the second variational property holds, which states  $[R_k^{k-1}] \equiv [P_{k-1}^k]^T$ . Variational coarsening then also produces the Schur complement.

For a one-dimensional problem,  $[L^{22}]$  is diagonal and its inverse can easily be found. In general however, this matrix is not a diagonal matrix, and a set of equations has to be solved in order to obtain the coarse-grid matrix and intergrid transfer operator matrices. For this general case, the coarse-grid operator as well as the prolongation operator and restriction operator lose their sparse nature. To illustrate this, figure C.5 presents a single triangular parent element and its four child elements. The fine-scale equations read  $[L^k][u^k] = [f^k]$ , and the discrete fine-scale operator is given by

$$[L^k] = \begin{bmatrix} L_{11}^k & 0 & 0 & L_{14}^k & L_{15}^k & L_{16}^k \\ 0 & L_{22}^k & 0 & L_{24}^k & L_{25}^k & 0 \\ 0 & 0 & L_{33}^k & L_{34}^k & 0 & L_{36}^k \\ L_{41}^k & L_{42}^k & L_{43}^k & L_{44}^k & L_{45}^k & L_{46}^k \\ L_{51}^k & L_{52}^k & 0 & L_{54}^k & L_{55}^k & 0 \\ L_{61}^k & 0 & L_{63}^k & L_{64}^k & 0 & L_{66}^k \end{bmatrix}. \quad (\text{C.42})$$

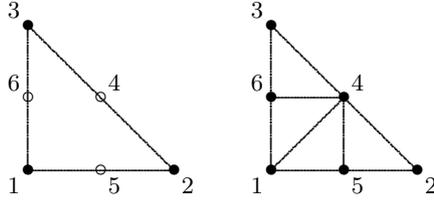


Figure C.5: Triangular parent and child elements.

The set of equations is decomposed as

$$\begin{bmatrix} [L^{11}] & [L^{12}] \\ [L^{21}] & [L^{22}] \end{bmatrix} \begin{bmatrix} [u^1] \\ [u^2] \end{bmatrix} = \begin{bmatrix} [f^1] \\ [f^2] \end{bmatrix}, \quad (\text{C.43})$$

where

$$[L^{22}] = \begin{bmatrix} L_{44}^k & L_{45}^k & L_{46}^k \\ L_{54}^k & L_{55}^k & 0 \\ L_{64}^k & 0 & L_{66}^k \end{bmatrix}. \quad (\text{C.44})$$

The inverse of this block matrix  $[A] = [L^{22}]^{-1}$  has a full structure, and the expression of its components gets complicated. The sub structuring technique builds super elements.

### Approximate Schur complement

To overcome the problem of inverting the matrix with f-node interactions, approximations of the discrete differential operator are used to obtain the coarse-grid operator. The general approach is to replace the set of equations by a modified set of equations, for which the inverse matrix can easily be obtained. The coarse-grid operator and intergrid transfer operator then follow from the Schur complement of the approximated fine-grid operator. It can be shown that the prolongation operator produces an optimal interpolation for the altered discretization, in the sense that variational coarsening of the approximated coarse-grid operator with this prolongation operator obtains the same result. However, the coarse-grid matrix does not in general maintain its sparse structure.

The modified set of algebraic equations  $[\tilde{L}^k] [u^k] = [\tilde{f}^k]$ , which follows from the approximation of the fine-grid discretization, will be written in block matrix-vector form as

$$\begin{bmatrix} [L^{11}] & [L^{12}] \\ [\tilde{L}^{21}] & [\tilde{L}^{22}] \end{bmatrix} \begin{bmatrix} [u^1] \\ [u^2] \end{bmatrix} = \begin{bmatrix} [f^1] \\ [\tilde{f}^2] \end{bmatrix}. \quad (\text{C.45})$$

For the approximation, firstly the computational costs for calculating the inverse of  $[\tilde{L}^{22}]$  should be low. Secondly, the product  $[\tilde{L}^{22}]^{-1} [\tilde{L}^{21}]$  should approximate  $[L^{22}]^{-1} [L^{21}]$ . Thirdly, the interpolation matrix has to be a local operator.

The Schur complement of the approximated matrix reads

$$[L^{k-1}] = [L^{11}] - [L^{12}] [\tilde{L}^{22}]^{-1} [\tilde{L}^{21}]. \quad (\text{C.46})$$

This results follows from variational coarsening as  $[L^{k-1}] = [\tilde{R}_{k-1}^k] [\tilde{L}^k] [\tilde{P}_{k-1}^k]$ , where the restriction matrix reads

$$[R_k^{k-1}] = \begin{bmatrix} [I] & -[L^{12}] [\tilde{L}^{22}]^{-1} \end{bmatrix}, \quad (\text{C.47})$$

and the prolongation matrix is formulated as

$$[\tilde{P}_{k-1}^k] = \begin{bmatrix} [I] \\ -[\tilde{L}^{22}]^{-1} [\tilde{L}^{21}] \end{bmatrix}. \quad (\text{C.48})$$

In general, the Galerkin approximation, where  $[R_k^{k-1}] \equiv [\tilde{P}_{k-1}^k]$ , does not obtain the same approximated Schur complement. Two modifications will be given to illustrate the approximation procedure for the model problem associated to figure C.5.

Reusken [103] presented the first modification. The equation for fine-grid node 5 reads

$$L_{51}u_1 + L_{52}u_2 + L_{54}u_4 + L_{55}u_5 = f_5. \quad (\text{C.49})$$

A linear interpolation for the unknown in fine-grid point 4:  $u_4 = (u_2 + u_3)/2$  then replaces the expression for node 5 by

$$L_{51}u_1 + \left(L_{52} + \frac{L_{54}}{2}\right)u_2 + \frac{L_{54}}{2}u_3 + L_{55}u_5 = f_5. \quad (\text{C.50})$$

Repeating the same procedure for unknowns  $u_4$  and  $u_6$  produces an approximated f-node to f-node connectivity matrix  $[\tilde{L}^{22}]$ , which reads

$$[\tilde{L}^{22}] = \begin{bmatrix} L_{44}^k & 0 & 0 \\ 0 & L_{55}^k & 0 \\ 0 & 0 & L_{66}^k \end{bmatrix}. \quad (\text{C.51})$$

The approximated f-node to c-node connectivity matrix is given by

$$[\tilde{L}^{21}] = \begin{bmatrix} L_{41}^k + \frac{L_{45}^k}{2} + \frac{L_{46}^k}{2} & L_{42}^k + \frac{L_{45}^k}{2} & L_{43}^k + \frac{L_{46}^k}{2} \\ L_{51}^k & L_{52}^k + \frac{L_{54}^k}{2} & \frac{L_{54}^k}{2} \\ L_{61}^k & \frac{L_{64}^k}{2} & L_{63}^k + \frac{L_{64}^k}{2} \end{bmatrix}. \quad (\text{C.52})$$

Now, the approximate Schur complement as well as the prolongation and restriction matrices can be calculated straight forwardly.

Wagner et al. [120] suggested the second modification. They proposed not to take a linear interpolation for the f-grid contributions but to take a matrix depended interpolation. The unknown in point 4 was pre-calculated as  $u_4 = (L_{42}u_2 + L_{43}u_3)/(L_{42} + L_{43})$ . Using this expression, the discrete equation for point 5 becomes

$$L_{51}u_1 + \left( L_{52} + \frac{L_{54}L_{42}}{L_{42} + L_{43}} \right) u_2 + \frac{L_{54}L_{43}}{L_{42} + L_{43}} u_3 + L_{55}u_4 = f_5 - \frac{L_{54}}{L_{44}} f_4. \quad (\text{C.53})$$

The coarse-scale operator follows from equation (C.46) and the prolongation operator is computed by equation (C.48). This equation produces an interpolation which is not dimensionally reduced. Wagner et al. applied their black box multigrid procedure successfully, solved an elliptic problem (steady state flow equation), a Poisson problem (transient flow equation) and a 'hyperbolic' problem (convection dominated transport equation). However, a disadvantage of both Wagner's and Reusken's technique is the fact that both introduce connectivities on the coarse scale, which are not present at a direct finite element coarse-scale computation.

### Variational coarsening over quadrilateral elements

Dendy [38] proposed directional lumping, based on a nine point finite difference molecule for regular grids. A directional lumping procedure reduces the interpolation dimension and generates a more local prolongation matrix. Variational coarsening uses this prolongation matrix, and produces a coarse-grid operator, which is based on the fine-grid discretization only.

Figure C.6 presents a coarse mesh and a fine mesh with coarse - intermediate - fine node ordering (c-i-f ordering). C-nodes denote nodes on the fine grid that are also

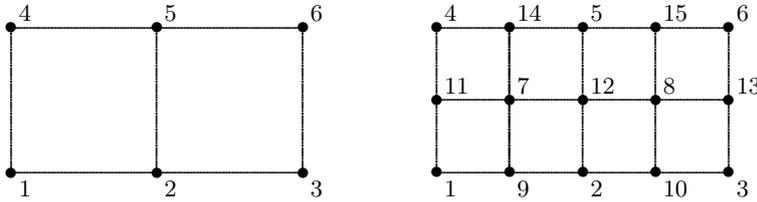


Figure C.6: Coarse mesh and c-i-f ordered fine mesh.

present in the coarse-grid  $\Omega^{k-1} \cap \Omega^k$ , (node 1, ..., 6). The points belonging only to the fine grid  $\Omega^k \setminus \Omega^{k-1}$ , will be called either f-nodes (node 9, ..., 15) or i-nodes (node 7 and 8). For a quadrilateral mesh, the i-nodes are interpolated from f-nodes and c-nodes, only when the f-node values are known. The approximation aims to relate f-nodes to c-nodes only. For this reason the fine-scale equation has to be approximated by

$$\begin{bmatrix} [L^{11}] & [L^{12}] & [L^{13}] \\ [L^{21}] & [L^{22}] & [L^{23}] \\ [\tilde{L}^{31}] & 0 & [\tilde{L}^{33}] \end{bmatrix} \begin{bmatrix} [u^1] \\ [u^2] \\ [u^3] \end{bmatrix} = \begin{bmatrix} [f^1] \\ [f^2] \\ [\tilde{f}^3] \end{bmatrix}, \quad (\text{C.54})$$

where the subvector  $[u^1]$  collects unknowns at the c-nodes,  $[u^2]$  contains unknowns at the i-nodes, and  $[u^3]$  collects unknowns at the f-nodes. In this approximation f-node to i-nodes connectivities, expressed by the block matrix  $[L^{32}]$ , are canceled out. The Schur complement of the approximated matrix reads

$$[L^{k-1}] = [L^{11}] - [ [L^{12}] \quad [L^{13}] ] \begin{bmatrix} [L^{22}] & [L^{23}] \\ 0 & [\tilde{L}^{33}] \end{bmatrix}^{-1} \begin{bmatrix} [L^{21}] \\ [\tilde{L}^{31}] \end{bmatrix}. \quad (\text{C.55})$$

Alternatively, the Schur complement can be obtained by variational coarsening that incorporates the Galerkin approximation. The lumping procedure will be explained for the mesh of quadrilateral elements shown in figure C.7. Here c-nodes  $1, \dots, 4$ , f-nodes  $4, \dots, 8$  and i-node 5 are the nodes on the fine mesh. The equation for f-nodes laying

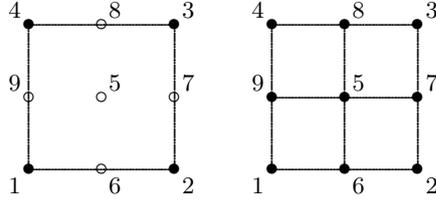


Figure C.7: Quadrilateral parent and child elements.

on a horizontal boundary is altered by vertical lumping. This will be demonstrated for node 6 in figure C.7. The algebraic equation for  $u_6$  reads

$$L_{61}u_1 + L_{62}u_2 + L_{65}u_5 + L_{66}u_6 + L_{67}u_7 + L_{69}u_9 = f_6, \quad (\text{C.56})$$

and the lumped expression is given by

$$(L_{65} + L_{66})u_6 + (L_{61} + L_{69})u_1 + (L_{62} + L_{67})u_2 = f_6. \quad (\text{C.57})$$

In the same way horizontal lumping is applied to modify the equations for f-nodes laying on a vertical boundary. The approximated f-node to f-node matrix reads

$$[\tilde{L}^{33}] = \begin{bmatrix} L_{66}^k + L_{65}^k & 0 & 0 & 0 \\ 0 & L_{77}^k + L_{75}^k & 0 & 0 \\ 0 & 0 & L_{88}^k + L_{85}^k & 0 \\ 0 & 0 & 0 & L_{99}^k + L_{95}^k \end{bmatrix}. \quad (\text{C.58})$$

The f-node to c-node matrix is approximated by

$$[\tilde{L}^{31}] = \begin{bmatrix} L_{61}^k + L_{69}^k & L_{62}^k + L_{67}^k & 0 & 0 \\ 0 & L_{72}^k + L_{76}^k & L_{73}^k + L_{78}^k & 0 \\ 0 & 0 & L_{83}^k + L_{87}^k & L_{84}^k + L_{89}^k \\ L_{91}^k + L_{96}^k & 0 & 0 & L_{94}^k + L_{98}^k \end{bmatrix}. \quad (\text{C.59})$$

The f-node to i-node matrix is replaced by the zero matrix. Matrix dependent prolongation operator coefficients for node 6 are found to be  $P_{61} = (L_{61} + L_{69})/(L_{66} + L_{65})$

and  $P_{62} = (L_{62} + L_{67})/(L_{66} + L_{65})$ . In the next step, f-node expressions are used to calculate the expression for i-node 5. The contribution of  $L_{56}u_6$  is replaced by  $L_{56}(L_{61} + L_{69})/(L_{66} + L_{65})u_1 + L_{56}(L_{62} + L_{67})/(L_{66} + L_{65})u_2$  and so on. The expression for node 5 is finally given by

$$\begin{aligned} & \left( L_{51} + L_{59} \frac{L_{91} + L_{96}}{L_{99} + L_{95}} + L_{56} \frac{L_{61} + L_{69}}{L_{66} + L_{65}} \right) u_1 \\ & + \left( L_{52} + L_{56} \frac{L_{62} + L_{67}}{L_{66} + L_{65}} + L_{57} \frac{L_{72} + L_{76}}{L_{77} + L_{75}} \right) u_2 \\ & + \left( L_{53} + L_{57} \frac{L_{73} + L_{78}}{L_{77} + L_{75}} + L_{58} \frac{L_{83} + L_{87}}{L_{88} + L_{85}} \right) u_3 \\ & + \left( L_{54} + L_{58} \frac{L_{84} + L_{89}}{L_{88} + L_{85}} + L_{59} \frac{L_{94} + L_{98}}{L_{99} + L_{95}} \right) u_4 + L_{55}u_5 = f_5. \quad (\text{C.60}) \end{aligned}$$

From this expression the prolongation operator coefficients are derived. Fine-grid unknowns follow from coarse-grid unknowns only, and interpolation does not have to be done in a prescribed order.

### Variational coarsening over triangular elements

This subsection proposes a new variational coarsening procedure over triangular elements. The procedure generalizes directional lumping over simplex elements, by adding basis functions and conserving the flux over an inner element boundary. Figure C.8 presents isolines of the basis functions  $N_2$  and  $N_3$ , and the newly constructed basis function  $\tilde{N}_1$ , which follows from  $\tilde{N}_1 = N_1 + N_4$ . These basis functions reduce the dimensions of the flow problem by directional lumping. The dimensionally reduced flow

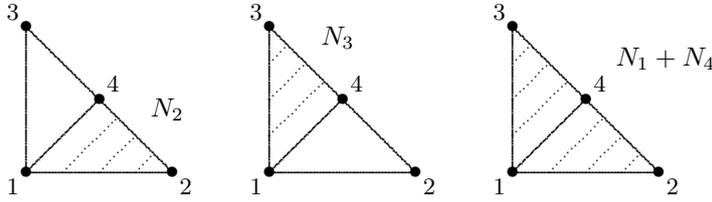


Figure C.8: Lumped fine element basis functions.

problem  $[\hat{L}^k] [\hat{u}^k] = [0]$  relates the pressure in node 4 (or 1) to the pressures in nodes 2 and 3. The associated connectivity matrix reads

$$[\tilde{L}^k] = \begin{bmatrix} \hat{L}_{11}^k & \hat{L}_{12}^k & \hat{L}_{13}^k \\ \hat{L}_{21}^k & \hat{L}_{22}^k & \hat{L}_{23}^k \\ \hat{L}_{31}^k & \hat{L}_{32}^k & \hat{L}_{33}^k \end{bmatrix}, \quad (\text{C.61})$$

where the components in the first row, are given by

$$\begin{aligned}\hat{L}_{11}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \left( \frac{\partial N_1^k}{\partial x_i} + \frac{\partial N_4^k}{\partial x_i} \right) \left( \frac{\partial N_1^k}{\partial x_j} + \frac{\partial N_4^k}{\partial x_j} \right) d\Omega \\ \hat{L}_{12}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \left( \frac{\partial N_1^k}{\partial x_i} + \frac{\partial N_4^k}{\partial x_i} \right) \frac{\partial N_2^k}{\partial x_j} d\Omega \\ \hat{L}_{13}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \left( \frac{\partial N_1^k}{\partial x_i} + \frac{\partial N_4^k}{\partial x_i} \right) \frac{\partial N_3^k}{\partial x_j} d\Omega,\end{aligned}$$

second row components read

$$\begin{aligned}\hat{L}_{21}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \frac{\partial N_2^k}{\partial x_i} \left( \frac{\partial N_1^k}{\partial x_j} + \frac{\partial N_4^k}{\partial x_j} \right) d\Omega \\ \hat{L}_{22}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \frac{\partial N_2^k}{\partial x_i} \frac{\partial N_2^k}{\partial x_j} d\Omega \\ \hat{L}_{23}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \frac{\partial N_2^k}{\partial x_i} \frac{\partial N_3^k}{\partial x_j} d\Omega,\end{aligned}$$

and third row components are written as

$$\begin{aligned}\hat{L}_{31}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \frac{\partial N_3^k}{\partial x_i} \left( \frac{\partial N_1^k}{\partial x_j} + \frac{\partial N_4^k}{\partial x_j} \right) d\Omega \\ \hat{L}_{32}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \frac{\partial N_3^k}{\partial x_i} \frac{\partial N_2^k}{\partial x_j} d\Omega \\ \hat{L}_{33}^{k-1} &= \int_{\Omega} K_{ij} \frac{k_r \rho}{\mu} \frac{\partial N_3^k}{\partial x_i} \frac{\partial N_3^k}{\partial x_j} d\Omega.\end{aligned}\tag{C.62}$$

Interpolation weights follow from the first row of the set of algebraic equations as

$$P_{42}^k = \hat{L}_{12}^{k-1} / \hat{L}_{11}^{k-1}, \quad P_{43}^k = \hat{L}_{13}^{k-1} / \hat{L}_{11}^{k-1},\tag{C.63}$$

or in components of the original matrix as

$$P_{42}^k = \frac{L_{12}^k + L_{42}^k}{L_{11}^k + L_{41}^k + L_{14}^k + L_{44}^k}, \quad P_{43}^k = \frac{L_{13}^k + L_{43}^k}{L_{11}^k + L_{41}^k + L_{14}^k + L_{44}^k}.\tag{C.64}$$

For a homogeneous element standard interpolation weights are found as  $P_{42}^k = 1/2$  and  $P_{43}^k = 1/2$ . In general, the modified basis function  $\tilde{N}_1$  gives an expression for the continuity of flow across the line which divides the coarse element into two child elements. Variational coarsening produces the coarse scale conductivity matrices as

$$[L^{k-1}] = [R_k^{k-1}] [L^k] [P_{k-1}^k].\tag{C.65}$$

A multistep refinement procedure splits a single triangular coarse element into four child elements. For this reason the coarsening procedure is carried out in two steps.

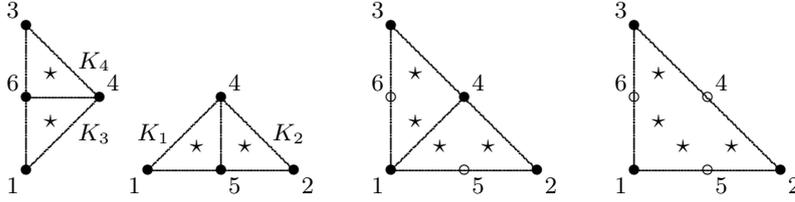


Figure C.9: Multistep coarsening of triangular elements.

Figure C.9 illustrates this sequential coarsening procedure. In the first step, element (6, 4, 3) and element (1, 4, 6) are mapped on the intermediate element (1, 4, 3). Element (5, 2, 4) and (1, 5, 4) produce the second intermediate element (1, 2, 4). In the next step, element (1, 4, 3) and (1, 2, 4) merge into the coarse-grid element (1, 2, 3).

The assemble of element (6, 4, 3) and (1, 4, 6) produce a local matrix for which the column and row numbers correspond to 1, 3, 4, 6. Variational coarsening maps this conductivity matrix to a new conductivity matrix with global numbering 1, 3, 4, which corresponds to the intermediate element (1, 4, 3). This variational coarsening step reads

$$\begin{bmatrix} L_{11}^{k-1} & L_{12}^{k-1} & L_{13}^{k-1} \\ L_{21}^{k-1} & L_{22}^{k-1} & L_{23}^{k-1} \\ L_{31}^{k-1} & L_{32}^{k-1} & L_{33}^{k-1} \end{bmatrix}_{lc} = \begin{bmatrix} 1 & 0 & 0 & R_{14} \\ 0 & 1 & 0 & R_{24} \\ 0 & 0 & 1 & 0 \end{bmatrix}_{lc} \begin{bmatrix} L_{11}^k & 0 & L_{13}^k & L_{14}^k \\ 0 & L_{22}^k & L_{23}^k & L_{24}^k \\ L_{31}^k & L_{32}^k & L_{33}^k & L_{34}^k \\ L_{41}^k & L_{42}^k & L_{43}^k & L_{44}^k \end{bmatrix}_{lc} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ P_{41} & P_{42} & 0 \end{bmatrix}_{lc}. \quad (\text{C.66})$$

If  $K_{xx} = K_{yy}$  and  $K_{xy} = K_{yx} = 0$ , then the conductivity matrix, which relates to element (1, 3, 4), and follows from variational coarsening and  $R_{16} = P_{61} = K_3/(K_3 + K_4)$ ,  $R_{36} = P_{63} = K_4/(K_3 + K_4)$ , reads

$$[L^{k*}]^{3,4} = \frac{\rho k_r}{2\mu} \begin{bmatrix} K_3 & 0 & -K_3 \\ 0 & K_4 & -K_4 \\ -K_3 & -K_4 & K_3 + K_4 \end{bmatrix}. \quad (\text{C.67})$$

These findings will be interpreted by constructing the coarse grid matrix for equivalent permeabilities. The coarse grid analogy for element (1, 3, 4) reads

$$\begin{aligned} [L^{k*}]^{3,4} &= \frac{\rho k_r}{4\mu} \begin{bmatrix} \langle K_{xx} \rangle + \langle K_{yy} \rangle & \langle K_{xx} \rangle - \langle K_{yy} \rangle & -2 \langle K_{xx} \rangle \\ \langle K_{xx} \rangle - \langle K_{yy} \rangle & \langle K_{xx} \rangle + \langle K_{yy} \rangle & -2 \langle K_{xx} \rangle \\ -2 \langle K_{xx} \rangle & -2 \langle K_{xx} \rangle & 4 \langle K_{xx} \rangle \end{bmatrix} \\ &+ \frac{\rho k_r}{4\mu} \begin{bmatrix} K_{xy}^v + K_{yx}^v & -K_{xy}^v + K_{yx}^v & -2K_{yx}^v \\ K_{xy}^v - K_{yx}^v & -K_{xy}^v - K_{yx}^v & 2K_{yx}^v \\ -2K_{xy}^v & 2K_{xy}^v & 0 \end{bmatrix}. \quad (\text{C.68}) \end{aligned}$$

For this case the averaged permeability coefficients read  $\langle K_{xx} \rangle = \langle K_{yy} \rangle = (K_3 + K_4)/2$  and  $\langle K_{xy} \rangle = \langle K_{yx} \rangle = (K_3 - K_4)/2$ .

For element (1, 2, 4), the conductivity matrix follows from variational coarsening and  $R_{15} = P_{51} = K_1/(K_1 + K_2)$ ,  $R_{25} = P_{52} = K_2/(K_1 + K_2)$ . This matrix reads

$$[L^{k*}]^{1,2} = \frac{\rho k_r}{2\mu} \begin{bmatrix} K_1 & 0 & -K_1 \\ 0 & K_2 & -K_2 \\ -K_1 & -K_2 & K_1 + K_2 \end{bmatrix}. \quad (\text{C.69})$$

The coarse grid analogy for element (1, 2, 4) reads

$$[L^{k*}]^{1,2} = \frac{\rho k_r}{4\mu} \begin{bmatrix} \langle K_{xx} \rangle + \langle K_{yy} \rangle & -\langle K_{xx} \rangle + \langle K_{yy} \rangle & -2 \langle K_{yy} \rangle \\ -\langle K_{xx} \rangle + \langle K_{yy} \rangle & \langle K_{xx} \rangle + \langle K_{yy} \rangle & -2 \langle K_{yy} \rangle \\ -2 \langle K_{yy} \rangle & -2 \langle K_{yy} \rangle & 4 \langle K_{yy} \rangle \end{bmatrix} + \frac{\rho k_r}{4\mu} \begin{bmatrix} K_{xy}^v + K_{yx}^v & K_{xy}^v - K_{yx}^v & -2K_{xy}^v \\ -K_{xy}^v + K_{yx}^v & -K_{xy}^v - K_{yx}^v & 2K_{yx}^v \\ -2K_{yx}^v & 2K_{yx}^v & 0 \end{bmatrix}. \quad (\text{C.70})$$

Here the averaged components read  $\langle K_{xx} \rangle = \langle K_{yy} \rangle = (K_1 + K_2)/2$  and  $\langle K_{xy} \rangle = \langle K_{yx} \rangle = (K_1 - K_2)/2$ . Components of the conductivity matrix for element (1, 2, 3) follow from both intermediate conductivity matrices, variational coarsening and setting  $R_{24} = P_{42} = K_2/(K_2 + K_4)$ ,  $R_{34} = P_{43} = K_4/(K_2 + K_4)$ .

The lumping process is repeated over the second step, where element (1, 4, 3) and element (1, 2, 4) merge into the coarse-grid element (1, 2, 3). The equivalent coefficients finally read

$$\begin{aligned} K_{xx}^v &= \frac{K_2 K_4}{2(K_2 + K_4)} + \frac{K_2 K_2 (K_1 + K_3)}{2(K_2 + K_4)(K_2 + K_4)}, \\ K_{yy}^v &= \frac{K_2 K_4}{2(K_2 + K_4)} + \frac{K_4 K_4 (K_1 + K_3)}{2(K_2 + K_4)(K_2 + K_4)}, \\ K_{xy}^v = K_{yx}^v &= -\frac{K_2 K_4}{2(K_2 + K_4)} + \frac{K_2 K_4 (K_1 + K_3)}{2(K_2 + K_4)(K_2 + K_4)}. \end{aligned} \quad (\text{C.71})$$

The method implicitly provides a symmetric permeability tensor and symmetric coarse scale matrix, but reduces the flow dimension in three directions. Due to the sequential setting this has a large effect on the resulting equivalent permeability as will be illustrated by the next test problem. Figure C.10 shows a model problem for which the hierarchical method fails. For element (1,2,4) the equivalent permeability  $K_{yy}^v$  follows from equation (C.71), and reads

$$K_{yy}^{v1} = \frac{K_a (K_a K_b + 3K_a K_a)}{2(K_a + K_b)(K_a + K_b)}. \quad (\text{C.72})$$

Whereas for element (1,3,4) this reads

$$K_{yy}^{v2} = \frac{K_a (K_a K_b + 3K_b K_b)}{2(K_a + K_b)(K_a + K_b)}. \quad (\text{C.73})$$

The interpolation weights follow from  $P_{42} = K_{yy}^{v1}/(K_{yy}^{v1} + K_{yy}^{v1})$  and  $P_{43} = K_{yy}^{v2}/(K_{yy}^{v1} + K_{yy}^{v1})$ . Only if the vertical permeability of element (1, 2, 4) and element (1, 3, 4) have

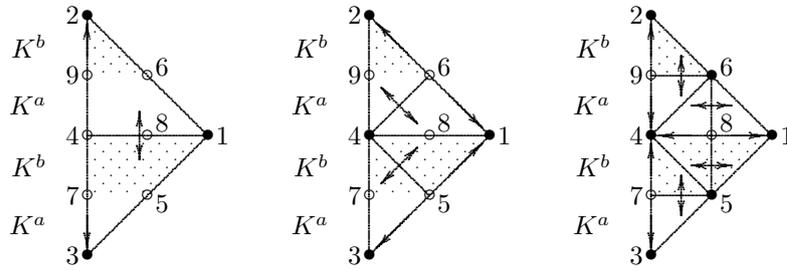


Figure C.10: Multi-step test problem.

the same value, then  $P_{42} = P_{43} = 0.5$ , as one would expect from a reduced flow computation over the element boundary. For this reason the proposed variational coarsening procedure over triangular elements has been rejected, and patches of simplex elements were proposed in this thesis.

# Notations

## Quantities

$a, A$	scalar
$\mathbf{v}$	vector
$\mathbf{A}$	second order tensor
$\mathbf{I}$	second order unit tensor
$v_a$	Cartesian vector components
$A_{ab}$	Cartesian tensor components
$[b]$	column
$[A]$	matrix
$[A^{ab}]$	block matrix
$\delta_{ab}$	Kronecker delta

## Operations

$\mathcal{L}$	linear operator
$\mathcal{N}$	nonlinear operator
$\Delta$	increment
$\nabla$	gradient operator
$\partial/\partial t$	partial time derivative
$D/Dt$	material time derivative
$\langle \cdot \rangle$	volume average operator
$ a $	absolute value of a scalar
$\mathbf{u} \cdot \mathbf{v}$	in-product or dot-product of two vectors
$\mathbf{u} * \mathbf{v}$	outer-product of two vectors
$\mathbf{u} \mathbf{v}$	dyadic-product or tensor-product of two vectors
$\ \cdot\ $	vector norm or tensor norm
$\mathbf{A}^D$	deviatoric part of a second-order tensor
$\mathbf{A}^T$	conjugation or transpose of a second-order tensor
$\mathbf{A}^{-1}$	inverse of a second-order tensor
$\mathbf{A}:\mathbf{B}$	scalar-product or double-dot product of two second-order tensors
$tr(\mathbf{A})$	trace or first invariant of a second-order tensor
$det(\mathbf{A})$	determinant or third invariant of a second-order tensor
$[A]^T$	transpose of a matrix

$[A]^{-1}$  inversion of a matrix

### Subscripts and Superscripts

$a, b$  matrix index  
 $(c)$  loading case  
 $e$  element  
 $g$  gas phase  
 $i, j$  space index  
 $k$  scale index  
 $l$  liquid or fluid phase  
 $n$  time step  
 $m$  gauss point  
 $r$  iteration step  
 $s$  solid phase  
 $x$  global coordinate  
 $\alpha$  phase indicator  
 $\xi$  local coordinate

### Integers

$n_b$  number of boundary elements  
 $n_d$  flow domain dimension  
 $n_e$  number of elements  
 $n_{en}$  number of element nodes  
 $n_k$  number of species  
 $n_n$  number of nodes  
 $n_u$  number of unknowns  
 $n^\alpha$  number of phases  
 $n_k^\alpha$  number of species per phase

### Units

[l] length  
[m] mass  
[t] time  
[T] temperature

### Symbols

$[A^p]$  flow stiffness matrix [ $l^{n_d-2}t^2$ ]  
 $[A^\omega]$  solute stiffness matrix [ $ml^{n_d-3}$ ]  
 $[A^T]$  heat stiffness matrix [ $ml^{n_d-1}t^{-2}T^{-1}$ ]  
 $b_l$  pore-connectivity parameter [-]  
 $b_n$  pore size distribution index [-]

$b_\alpha$	Brooks-Corey curve-fitting parameter [ $l^{-1}$ ]
$b_\kappa$	Brooks-Corey curve-fitting parameter [ $-$ ]
$[B^p]$	flow force vector [ $ml^{n_d-3}$ ]
$[B^\omega]$	solute force vector [ $ml^{n_d-3}$ ]
$[B^T]$	heat force vector [ $ml^{n_d-1}t^{-2}$ ]
$c$	specific heat capacity [ $l^2t^{-2}T^{-1}$ ]
$C_f$	Forchheimer coefficient [ $l^{-1}t$ ]
$C_h$	non-Fickian high concentration dispersion coefficient [ $m^{-1}l^2t$ ]
$Cr$	Courant number [ $-$ ]
$D_c$	coefficient of molecular dispersion [ $l^2t^{-1}$ ]
$\mathbf{D}$	hydrodynamic dispersion tensor [ $l^2t^{-1}$ ]
$\mathbf{D}_m$	mechanical dispersion tensor [ $l^2t^{-1}$ ]
$E_h$	internal thermal energy density [ $l^2t^{-2}$ ]
$\mathbf{f}_k$	internal drag vector or momentum exchange [ $ml^{-2}t^{-2}$ ]
$[F^p]$	flow force vector [ $ml^{n_d-3}t^{-1}$ ]
$[F^\omega]$	solute force vector [ $ml^{n_d-3}t^{-1}$ ]
$[F^T]$	heat force vector [ $ml^{n_d-1}t^{-3}$ ]
$g_a$	Genuchten shape factor [ $l^{-1}$ ]
$g_l, g_m, g_n$	Genuchten shape factors [ $-$ ]
$\mathbf{g}$	gravitational body force vector [ $lt^{-2}$ ]
$H$	energy source term for heat supply [ $ml^{-1}t^{-3}$ ]
$\mathbf{H}$	hydrodynamic thermo-dispersion tensor [ $mlt^{-3}T^{-1}$ ]
$j$	magnitude of the non-convective mass flux vector [ $ml^{-2}t^{-1}$ ]
$\mathbf{j}$	non-convective (dispersive and diffusive) flux vector [ $ml^{-2}t^{-1}$ ]
$\mathbf{j}_h$	non-convective heat energy flux [ $mt^{-3}$ ]
$k_r$	relative permeability [ $-$ ]
$M$	mass source term [ $ml^{-3}t^{-1}$ ]
$[M^p]$	flow compressibility matrix [ $l^{n_d-2}t^2$ ]
$[M^\omega]$	solute compressibility matrix [ $ml^{n_d-3}$ ]
$[M^T]$	heat compressibility matrix [ $ml^{n_d-1}t^{-2}T^{-1}$ ]
$n$	porosity [ $-$ ]
$\mathbf{n}$	unit normal vector [ $-$ ]
$p$	pressure [ $ml^{-1}t^{-2}$ ]
$p_c$	capillary pressure [ $ml^{-1}t^{-2}$ ]
$p_e$	evaporation threshold pressure [ $ml^{-1}t^{-2}$ ]
$p_p$	ponding threshold pressure [ $ml^{-1}t^{-2}$ ]
$Pe$	Peclet number [ $-$ ]
$q$	magnitude of the volumetric flux density [ $lt^{-1}$ ]
$q_n$	boundary source flux [ $lt^{-1}$ ]
$\mathbf{q}$	volumetric flux density [ $lt^{-1}$ ]
$Q$	bulk volumetric solute source term [ $t^{-1}$ ]
$r$	resistance term [ $l^{-1}t$ ]
$R$	chemical reaction rate term [ $ml^{-3}t^{-1}$ ]
$Ra$	Rayleigh number [ $-$ ]
$S$	saturation of the liquid phase in the void space [ $-$ ]

$S_r$	minimal saturation [-]
$S_s$	maximum saturated [-]
$[S^p]$	flow conductivity matrix [ $l^{n_d-2}t$ ]
$[S^\omega]$	solute conductivity matrix [ $ml^{n_d-3}t^{-1}$ ]
$[S^T]$	heat conductivity matrix [ $ml^{n_d-1}t^{-3}T^{-1}$ ]
$t$	time [t]
$\mathbf{T}$	tortuosity tensor [-]
$T$	temperature [T]
$\mathbf{u}$	displacement vector [l]
$\mathbf{v}$	velocity vector [ $lt^{-1}$ ]
$\mathbf{x}$	spatial coordinate vector [l]
$w$	Gauss weight [-]
$x, y, z$	global coordinates [l]
$\alpha_l$	longitudinal dispersivity [l]
$\alpha_p$	compressibility of the solid skeleton [ $m^{-1}lt^2$ ]
$\alpha_t$	transverse dispersivity [l]
$\beta_{p^l}$	liquid phase compressibility [ $m^{-1}lt^2$ ]
$\beta_{\omega_k^l}$	volumetric solute expansion coefficient [-]
$\beta_{T^l}$	thermal expansion coefficient [ $T^{-1}$ ]
$\Gamma$	domain boundary [ $l^{n_d-1}$ ]
$\gamma$	volumetric body force [ $ml^{-2}t^{-2}$ ]
$\boldsymbol{\varepsilon}$	linear strain tensor [-]
$\varepsilon_v$	volumetric strain [-]
$\theta$	time weighing factor [-]
$\mathbf{K}$	intrinsic permeability tensor [ $l^2$ ]
$\mathbf{K}_d$	hydraulic conductivity tensor [ $lt^{-1}$ ]
$\lambda$	thermal conductivity [ $mlt^{-3}T^{-1}$ ]
$\lambda_l$	Lamé constant [ $ml^{-1}t^{-2}$ ]
$\mu$	dynamic viscosity [ $ml^{-1}t^{-1}$ ]
$\mu_k$	chemical potential of species $k$ [ $l^2t^{-2}$ ]
$\mu_l$	Lamé constant [ $ml^{-1}t^{-2}$ ]
$\nu$	specific volume [ $m^{-1}l^3$ ]
$\xi, \eta, \zeta$	local coordinates [l]
$\rho$	density [ $ml^{-3}$ ]
$\boldsymbol{\sigma}_k$	viscous stress tensor [ $ml^{-1}t^{-2}$ ]
$\boldsymbol{\sigma}$	Cauchy (total) stress tensor [ $ml^{-1}t^{-2}$ ]
$\boldsymbol{\sigma}'$	effective stress tensor [ $ml^{-1}t^{-2}$ ]
$\boldsymbol{\tau}$	boundary traction tensor [ $mt^{-2}$ ]
$[\Upsilon]$	pressure-dissipation matrix [ $m^2l^{-4}t^{-3}$ ]
$\phi$	volume fraction [-]
$\psi$	pressure head [l]
$\psi_a$	air-entry pressure head [l]
$\omega$	mass fraction [-]
$\Omega$	domain [ $l^{n_d}$ ]

# Bibliography

- [1] J. E. Aarnes, V. L. Hauge, and Y. Efendiev. Coarsening of three-dimensional structured and unstructured grids for subsurface flow. *Advances in Water Resources*, 30(11):2177–2193, November 2007.
- [2] J.E. Aarnes, V. Kippe, and K.-A. Lie. Mixed multiscale finite elements and streamline methods for reservoir simulation of large geo-models. *Advances in Water Resources*, 28:257–271, 2005.
- [3] J.E. Aarnes, S. Krogstad, and K.-A. Lie. Multiscale mixed/mimetic methods on corner-point grids. *Computational Geosciences*, 12(3):297–315, September 2008.
- [4] PH. Ackerer, A. Younes, and R. Mose. Modeling variable density flow and solute transport in porous medium: 1. numerical model and verification. *Transport in Porous Media*, 35:345–373, 1999.
- [5] R.E. Alcouffe, A. Brandt, J.E. Dendy, Jr., and J.W. Painter. The multigrid method for the diffusion equation with strongly discontinuous coefficients. *SIAM Journal on Scientific Computing*, 2(4):430–453, 1981.
- [6] T. Arbogast. The implementation of a locally conservative numerical subgrid upscaling scheme for two-phase Darcy flow. *Computational Geosciences*, 6:453–481, 2002.
- [7] T. Arbogast and S.L. Bryant. Numerical subgrid upscaling for waterflood simulations. In *16th SPE Symposium on Reservoir Simulation*, page SPE66375, Houston, Texas, 2001. Society of Petroleum Engineers.
- [8] J.-L. Auriault and J. Lewandowska. Upscaling: Cell symmetries and scale separation. *Transport in Porous Media*, 43:473–485, 2001.
- [9] K. Aziz and A. Settari. *Petroleum Reservoir Simulation*. Blitzprint Ltd., Calgary, Alberta, 2002.
- [10] K.-J. Bathe. *Finite Element Procedures*. Prentice-Hall, 1996.
- [11] M. Bause and P. Knabner. Computation of variably saturated flow by adaptive mixed hybrid finite element methods. *Advances in Water Resources*, 27:565–581, 2004.

- [12] J. Bear. *Dynamics of Fluids in Porous Media*. Dover Publications, 1988.
- [13] J. Bear and A. Verruijt. *Modeling Groundwater Flow and Pollution*, volume 2 of *Theory and Applications of Transport in Porous Media*. D. Reidel Publishing Company, Dordrecht, The Netherlands, 1987.
- [14] T. Belytschko, W.K. Liu, and B. Moran. *Nonlinear Finite Elements for Continua and Structures*. Wiley, USA, 2008.
- [15] R.B. Bird, W.E. Stewart, and E.N. Lightfoot. *Transport Phenomena*. John Wiley and Sons, 1960.
- [16] Ø. Bø e. Analysis of an upscaling method based on conservation of dissipation. *Transport in Porous Media*, 17:77–86, 1994.
- [17] A.G.L. Borthwick, R.D. Marchant, and G.J.M. Copeland. Adaptive hierarchical grid model of water-born pollutant dispersion. *Advances in Water Resources*, 23:849–865, 2000.
- [18] W.L. Briggs, H. van Emden, and S.F. McCormick. *A Multigrid Tutorial*. SIAM, Society for Industrial and Applied Mathematics, USA, 2000.
- [19] R.B.J. Brinkgreve. *Plaxis, Finite Element Code for Soil and Rock Analyses*. Plaxis B.V., Delft, The Netherlands, 8th edition, 2002.
- [20] R.B.J. Brinkgreve, R. Al-Khoury, and J.M. van Esch. *Plaxflow*. Plaxis B.V., Delft, The Netherlands, 1th edition, 2003.
- [21] M.A. Bùes and C. Oltean. Numerical simulations for saltwater intrusion by the mixed hybrid finite element method and discontinuous finite element method. *Transport in Porous Media*, 40:171–200, 2000.
- [22] F. Chen and L. Ren. Application of the finite difference heterogeneous multiscale method to the Richards’ equation. *Water Resources Research*, 44:W07413, July 2008.
- [23] T. Chen, M.G. Gerritsen, J.V. Lambers, and L.J. Durlofsky. Global variable compact multipoint methods for accurate upscaling with full-tensor effects. *Computational Geosciences*, 14(1), January 2010.
- [24] Z. Chen and T.Y. Hou. A mixed multi scale finite element method for elliptic problems with oscillating coefficients. *Mathematics of Computation*, 72(242):541–576, 2002.
- [25] H.-P. Cheng, G.-T. Yeh, J. Xu, M.-H Li, and R. Carsel. A study of incorporating the multigrid method into the three-dimensional finite element discretization: A modular setting and application. *International Journal for Numerical Methods in Engineering*, 41:499–526, 1998.

- [26] M.A. Christie and M.J. Blunt. Tenth spe comparative solution project: A comparison of upscaling techniques. *SPE Reservoir Evaluation & Engineering*, pages 308–317, August 2001.
- [27] J. Chu, Y. Efendiev, V. Ginting, and T.Y. Hou. Flow based oversampling technique for multiscale finite element methods. *Advances in Water Resources*, 31(4):599–608, April 2008.
- [28] C.R. Cole and H.P. Foote. *Multigrid Methods for Solving Multiscale Transport Problems*, pages 273–303. Volume 10 of Cushman [29], 1990.
- [29] J.H. Cushman. *Dynamics of Fluids in Hierarchical Porous Media*. Academic Press, 1990.
- [30] J.H. Cushman, L.S. Bennethum, and B.X. Hu. A primer on upscaling tools for porous media. *Advances in Water Resources*, 25:1043–1067, 2002.
- [31] G. Dagan. *Flow and Transport in Porous Formations*. Springer, 1989.
- [32] G. de Josselin de Jong. Longitudinal and transverse diffusion in granular deposits. *Transactions, American Geophysical Union*, 39:67–74, 1958.
- [33] G. de Marsily. *Quantitative Hydrogeology, Groundwater Hydrology for Engineers*. Academic Press, 1986.
- [34] P.M. de Zeeuw. Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. *Journal of Computational Mathematics*, 33:1–27, 1990.
- [35] L. Demkowicz. *Computing with hp-Adaptive Finite Elements, Frontiers: Three Dimensional Elliptic and Maxwell Problems*, volume 2. Chapman & Hall/CRC, Boca Raton, USA, 2007.
- [36] L. Demkowicz. *Computing with hp-Adaptive Finite Elements, One and Two Dimensional Elliptic and Maxwell Problems*, volume 1. Chapman & Hall/CRC, Boca Raton, USA, 2007.
- [37] J.E. Dendy, Jr. Black box multigrid. *Journal of Computational Physics*, 48:366–386, 1982.
- [38] J.E. Dendy, Jr. Black box multigrid for non-symmetric problems. *Applied Mathematics and Computation*, 13:261–283, 1983.
- [39] H.-J.G. Diersch. *Feflow, Finite Element Subsurface Flow & Transport Simulation System, Users Manual*. Wasy Institute for Water Resources Planning and System Research, Berlin, Germany, 2005.
- [40] H.-J.G. Diersch. *Feflow, Finite Element Subsurface Flow & Transport Simulation System, White papers*. Wasy Institute for Water Resources Planning and System Research, Berlin, Germany, 2005.

- [41] H.-J.G Diersch and O. Kolditz. Variable-density flow and transport in porous media: Approaches and challenges. *Advances in Water Resources*, 25:899–944, 2002.
- [42] L.J. Durlofsky. Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. *Water Resources Research*, 27(5):699–708, 1991.
- [43] L.J. Durlofsky. Representation of grid block permeability in coarse scale models of randomly heterogeneous porous media. *Water Resources Research*, 28(7):1791–1800, 1992.
- [44] L.J. Durlofsky, Y. Efendiev, and V. Ginting. An adaptive local/global multiscale finite volume element method for two-phase flow simulations. *Advances in Water Resources*, 30(3):576–588, March 2007.
- [45] M.G. Edwards and C.F. Roger. Multigrid and renormalization for reservoir simulation. In P.W. Hemker and P. Wesseling, editors, *Multigrid methods IV*, pages 189–200, Amsterdam, 1994. Birkhäuser Verlag.
- [46] Y. Efendiev, T. Hou, and T. Strinopoulos. Multiscale simulations of porous media flows in flow-based coordinate system. *Computational Geosciences*, 12(3):257–272, September 2008.
- [47] J.W. Elder. Numerical experiments with free convection in a vertical slot. *Journal of Fluid Mechanics*, 24(4):823–843, 1966.
- [48] H.I. Ene and D. Poliševski. *Thermal Flow in Porous Media*, volume 1 of *Theory and Applications of Transport in Porous Media*. D. Reidel Publishing Company, 1987.
- [49] M.W. Farthing and C.T. Miller. A comparison of high-resolution, finite-volume, adaptive-stencil schemes for simulating advective-dispersive transport. *Advances in Water Resources*, 24:29–48, 2001.
- [50] P.A. Forsyth, Y.S. Wu, and K. Pruess. Robust numerical methods for saturated-unsaturated flow with dry initial conditions in heterogeneous media. *Advances in Water Resources*, 18:25–38, 1995.
- [51] D.G. Fredlund and H. Rahardjo. *Soil Mechanics for Unsaturated Soils*. John Wiley & Sons, 1993.
- [52] R.A. Freeze and J.A. Cherry. *Groundwater*. Prentice Hall, 1979.
- [53] G. Galeati, G. Gambolati, and S.P. Neuman. Coupled and partially coupled Euler-Lagrangian model of freshwater-seawater mixing. *Water Resources Research*, 28(1):149–165, 1992.
- [54] Y. Gautier and B. Noetinger. Preferential flow-paths detection for heterogeneous reservoirs using a new renormalization technique. In P.R. Thomassen, editor, *4th European Conference on the Mathematics of Oil Recovery*, 1992.

- [55] L.W. Gelhar. *Stochastic Subsurface Hydrology*. Prentice-Hall, 1993.
- [56] M. Gerritsen and J.V. Lambers. Integration of local/global upscaling and grid adaptivity for simulation of subsurface flow in heterogeneous formations. *Computational Geosciences*, 12(2), June 2008.
- [57] G.H. Golub and C.F. van Loan. *Matrix Computations*. John Hopkins University Press, London, United Kingdom, 3rd edition, 1996.
- [58] H. Gotovac, V. Cvetkovic, and R. Andricevic. Adaptive fup multi-resolution approach to flow and advective transport in highly heterogeneous porous media: Methodology, accuracy and convergence. *Advances in Water Resources*, 32(6):779–968, June 2009.
- [59] M. Griebel and S. Knappek. A multigrid-homogenization method. *13th GAMM-Seminar Kiel on Numerical Treatment of Multi Scale Problems*, pages 187–202, 1997.
- [60] W. Hackbush. *Multi-Grid Methods and Applications*, volume 4 of *Springer Series in Computational Mathematics*. Springer, 2003.
- [61] S.M. Hassanizadeh. On the transient non-Fickian dispersion theory. *Transport in Porous Media*, 23:107–124, 1996.
- [62] S.M. Hassanizadeh and A. Leijnse. A non-linear theory of high-concentration-gradient dispersion in porous media. *Advances in Water Resources*, 18(4):203–215, 1995.
- [63] X. He and L. Ren. Finite volume multiscale finite element method for solving the groundwater flow problems in heterogeneous porous media. *Water Resources Research*, 41:W10417, 2005.
- [64] K.J. Hersvik and M.S. Espedal. Adaptive hierarchical upscaling of flow in heterogeneous reservoirs based on an a posteriori error estimate. *Computational Geosciences*, 2:311–336, 1998.
- [65] L. Holden and B.F. Nielsen. Global upscaling of permeability in heterogeneous reservoirs; the output least squares (OLS) method. *Transport in Porous Media*, 40:115–143, 2000.
- [66] E. Holzbecher. *Modeling Density-Driven Flow in Porous Media*. Springer, 1998.
- [67] U. Hornung. *Homogenization and Porous Media*. Springer, 1997.
- [68] T.Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of Computational Physics*, 134:169–189, 1997.
- [69] T.Y. Hou, X.-H. Wu, and Z. Cai. Convergence of the multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Mathematics of Computation*, 68(227):913–943, 1999.

- [70] T.J.R. Hughes. *The Finite Element Method, Linear Static and Dynamic Finite Element Analysis*. Dover Publications, Mineola, New York, USA, 2000.
- [71] P.S. Huyakorn and G.F. Pinder. *Computational Methods in Subsurface Flow*. Academic Press, 1983.
- [72] J. Istok. *Groundwater Modeling by the Finite Element Method*, volume 13 of *Water Resources Monograph*. American Geophysical Union, 1990.
- [73] P. Jenny, S.H. Lee, and H.A. Tchelepi. Multi-scale finite-volume method for elliptic problems in subsurface flow simulation. *Journal of Computational Physics*, 187:47–67, 2003.
- [74] L. Jiang, Y. Efendiev, and I. Mishev. Mixed multiscale finite element methods using approximate global information based on partial upscaling. *Computational Geosciences*, 14(2), March 2010.
- [75] R. Juanes and F.X. Dub. A locally conservative variational multiscale method for the simulation of porous media flow with multiscale source terms. *Computational Geosciences*, 12(3):273–295, September 2008.
- [76] P.R. King. The use of renormalization for calculating effective permeability. *Transport in Porous Media*, 4:37–58, 1989.
- [77] P.R. King. Upscaling permeability: Error analysis for renormalization. *Transport in Porous Media*, 23:337–354, 1996.
- [78] K.L. Kipp, Jr. *HST3D, A Computer Code for Simulation of Heat and Solute Transport in Three-Dimensional Ground-Water Flow Systems*. U.S. Geological Survey, Denver, Colorado, USA, 1987.
- [79] K.L. Kipp, Jr. *Guide to the revisited Heat and Solute Transport Simulator: Hst3d - Version 2*. U.S. Geological Survey, Denver, Colorado, USA, 1997.
- [80] V. Kippe, J.E. Aarnes, and K.-A. Lie. Multiscale finite-element methods for elliptic problems in porous media flow. In *Proceedings of the 16th International Conference on Computational Methods in Water Resources*, Copenhagen, Denmark, 2006.
- [81] V. Kippe, J.E. Aarnes, and K.-A. Lie. A comparison of multiscale methods for elliptic problems in porous media flow. *Computational Geosciences*, 12(3):377–398, September 2008.
- [82] S. Knapek. Matrix-dependent multigrid homogenization for diffusion problems. *SIAM Journal on Scientific Computing*, 20(2):515–533, 1998.
- [83] O. Kolditz, R. Ratke, H.-J.G. Diersch, and W. Zielke. Coupled groundwater flow and transport: 1. verification of variable density flow and transport models. *Advances in Water Resources*, 21(1):27–46, 1998.

- [84] J.B. Kool and J.C. Parker. Development and evaluation of closed-form expressions for hysteretic soil hydraulic properties. *Water Resources Research*, 23(1):105–114, 1987.
- [85] W.M. Lai, D. Rubin, and E. Kreml. *Introduction to Continuum Mechanics*. Pergamon Press, 3rd edition, 1993.
- [86] J.V. Lambers, M.G. Gerritsen, and B.T. Mallison. Accurate local upscaling with variable compact multipoint transmissibility calculations. *Computational Geosciences*, 12(3), September 2008.
- [87] S.H. Lee, C. Wolfsteiner, and H.A. Tchelepi. Multiscale finite-volume formulation for multiphase flow in porous media: Black oil formulation of compressible, three-phase flow with gravity. *Computational Geosciences*, 12(3):351–366, September 2008.
- [88] A. Leijnse. *Three-Dimensional Modeling of Coupled Flow and Transport in Porous Media*. PhD thesis, University of Notre Dame, Notre Dame, Indiana, USA, 1992.
- [89] D.A. Lever and C.P. Jackson. On the equations for the flow of concentrated salt solution through a porous medium. *Harwell Laboratory report AERE-R 11765*, 1985.
- [90] M.-H. Li, H.-P. Cheng, and G.-T. Yeh. Solving 3d subsurface flow and transport with adaptive multigrid. *Journal of Hydrological Engineering*, pages 74–81, 2000.
- [91] I. Lunati and P. Jenny. Multiscale finite-volume method for density-driven flow in porous media. *Computational Geosciences*, 12(3):337–350, September 2008.
- [92] S.P. MacLachlan and J.D. Moulton. Multilevel upscaling through variational coarsening. *Water Resources Research*, 42:W02418, 2006.
- [93] S. Mehl and M.C. Hill. Three-dimensional local grid refinement for block-centered finite-difference groundwater models using iteratively coupled shared nodes: a new method of interpolation and analysis of errors. *Advances in Water resources*, 27:899–912, 2004.
- [94] C.C. Mei. Method of homogenization applied to dispersion in porous media. *Transport in Porous Media*, 9:261–274, 1992.
- [95] J.D. Moulton, J.E. Dendy, Jr., and J.M. Hyman. The black box multigrid numerical homogenization algorithm. *Journal of Computational Physics*, 142:80–108, 1998.
- [96] B.F. Nielsen and A. Tvieta. An upscaling method for one-phase flow in heterogeneous reservoirs. a weighted output least squares (WOLS) approach. *Computational Geosciences*, 2:93–123, 1998.

- [97] S. Nilsen and M.S. Espedal. Wavelet upscaling based of piecewise bilinear approximation on the permeability field. *Transport in Porous Media*, 23:125–134, 1996.
- [98] J. M. Nordbotten and P. E. Bjørstad. On the relationship between the multiscale finite-volume method and domain decomposition preconditioners. *Computational Geosciences*, 12(3):367–376, September 2008.
- [99] W.H. Press, W.T. Vetterling, S.A. Teukolsky, and B.P. Flannery. *Numerical recipes in C, The Art of Scientific Computing*. Cambridge University Press, New York, USA, 2nd edition, 2002.
- [100] M. Putti and C. Paniconi. Picard and Newton linearization for the coupled model of saltwater intrusion in aquifers. *Advances in Water Resources*, 18(3):159–170, 1995.
- [101] Ph. Renard and G. de Marsily. Calculating equivalent permeability: A review. *Advances in Water Resources*, 20(5–6):253–278, 1997.
- [102] A. Reusken. Multigrid with matrix-dependent transfer operators for a singular perturbation problem. *Computing*, 50:199–211, 1993.
- [103] A. Reusken. Multigrid with matrix-dependent transfer operators for convection-diffusion problems. In *Proceedings of the Fourth European Multigrid Conference*, Brinkhauser, Basel, 1994. INSM.
- [104] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, 2nd edition, 2000.
- [105] A.E. Scheidegger. General theory of dispersion in porous media. *Journal of Geophysical Research*, 66(10):3273–3278, 1961.
- [106] R.J. Schotting, H. Moser, and S.M. Hassanizadeh. High-concentration-gradient dispersion in porous media: Experiments, analysis and approximations. *Advances in Water Resources*, 22(7):665–680, 1999.
- [107] M.J. Simpson and T.P. Clement. Theoretical analysis of the worthiness of Henry and Elder problems as benchmarks of density-dependent groundwater flow models. *Advances in Water Resources*, 26:17–31, 2003.
- [108] C.W.M. Sitters. *Continuum Mechanics*. Delft University of Technology, Delft, The Netherlands, 1996.
- [109] J. Trangenstein. Multi-scale iterative techniques and adaptive mesh refinement for flow in porous media. *Advances in Water Resources*, 25:1175–1213, 2002.
- [110] U. Trottenberg, C. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, 2001.
- [111] J.C. van Dam. *Field-Scale Water Flow and Solute Transport, SWAP Model Concepts, Parameter Estimation and Case Studies*. PhD thesis, Wageningen Universiteit, Wageningen, The Netherlands, 2000.

- [112] J.M. van Esch. Adaptive grid modeling of unsaturated groundwater flow. In F.B.J. Barends and P.M.P.C. Steijger, editors, *Learned and Applied Soil Mechanics Out of Delft*, pages 9–14, Delft, The Netherlands, 2002. Delft University of Technology, A.A. Balkema.
- [113] J.M. van Esch. Adaptive multigrid modeling of density dependent groundwater flow. In R.H. Boekelman et al., editor, *Proceedings 17th Salt Water Intrusion Meeting*, pages 82–92, Delft, The Netherlands, 2002. Delft University of Technology, Hydrology & Ecology Section, Faculty of Civil Engineering and Geosciences.
- [114] J.M. van Esch, M.A. Van, R.F.A. Hendriks, and J.J.H. van den Akker. Geohydrologic design procedure for peat dykes under drying conditions. In G.N. Pande and S. Pietruszczak, editors, *Numerical Models in Geomechanics (Numog X)*. Taylor & Francis Group, 2007.
- [115] J.M. van Esch, A.F. van Tol, H.R. Havinga, A.M.W. Duijvenstijn, B.J. Schat, and J.C.W.M. de Wit. Functional analyses of jet grout bodies based on Monte Carlo simulations. In G. Barla and M. Barla, editors, *Proceedings of the 11th International Conference on Computer Methods and Advances in Geomechanics*, pages 643–650, Torino, Italy, 2005. Department of Structural and Geotechnical Engineering Politecnico di Torino, Patron Editore.
- [116] A. Verruijt. *Computational Geomechanics*, volume 7 of *Theory and Applications of Transport in Porous Media*. Kluwer Academic Publishers, 1995.
- [117] T. Vogel, M.Th. van Genuchten, and M. Cislerova. Effect of the shape of the soil hydraulic functions near saturation on variably-saturated flow predictions. *Advances in Water Resources*, 24:133–144, 2001.
- [118] C.I. Voss and A.M. Provost. *SUTRA, A Model for Saturated-Unsaturated, Variable-Density Ground-Water Flow with Solute or Energy Transport*. U.S. Geological Survey, Reston, Virginia, USA, 2003.
- [119] C.B. Vreugdenhil. *Computational Hydraulics*. Springer, 1989.
- [120] C. Wagner, W. Kinzelbach, and G. Wittum. Schur-complement multigrid, a robust method for groundwater flow and transport problems. *Numerische Mathematik*, 75:523–545, 1997.
- [121] D. Weatherill, C.T. Simmons, C.I. Voss, and N.I. Robinson. Testing density-dependent groundwater flow models: Two-dimensional steady state unstable convection in infinite, finite and inclined porous layers. *Advances in Water Resources*, 27:547–562, 2004.
- [122] P. Wesseling. *An Introduction to Multigrid Methods*. Edwards, Philadelphia, USA, 2004.
- [123] S. Whitaker. *The Method of Volume Averaging*, volume 13 of *Theory and Applications of Transport in Porous Media*. Kluwer Academic Publishers, 1999.

- 
- [124] S. Ye, Y. Xue, and C. Xie. Application of the multiscale finite element method to flow in heterogeneous porous media. *Water Resources Research*, 40:W09202, 2004.
- [125] G.-T. Yeh. *Computational Subsurface Hydrology, Fluid Flows*. Kluwer Academic Publishers, 1999.
- [126] G.T. Yeh. A Lagrangian-Eulerian method with zoomable hidden fine-mesh approach to solving advection-dispersion equations. *Water Resources Research*, 26(6):1133–1144, 1990.
- [127] O.C. Zienkiewicz and R.L. Taylor. *The Finite Element Method, Basic Formulation and Linear Problems*, volume 1. McGraw-Hill Book Company, 4th edition, 1989.
- [128] O.C. Zienkiewicz and R.L. Taylor. *The Finite Element Method, Solid and Fluid Mechanics Dynamics and Non-Linearity*, volume 2. McGraw-Hill Book Company, 4th edition, 1989.
- [129] W. Zijl and A. Trykozko. Numerical homogenization of the absolute permeability using the conformal-nodal and mixed-hybrid finite element method. *Transport in Porous Media*, 44:32–62, 2001.

# Samenvatting

Natuurlijke geologische formaties vertonen veelal een structurele en functionele heterogeniteit over vele ordes van grootte in de ruimte en de tijd. In geohydrologische simulaties wordt de structurele hiërarchie van fysische subeenheden met een geologisch model beschreven en de functionele hiërarchie van het stromingsproces met een stromingsmodel. Stromingseigenschappen zoals druk, flux en dissipatie kunnen aan elkaar gerelateerd worden met constitutieve relaties en parameters van de subeenheden zoals porositeit en hydraulische doorlatendheid. Componenten van de hydraulische doorlatendheid zijn de (stationaire) intrinsieke doorlatendheid en de (tijdsafhankelijke) relatieve doorlatendheid. De hydraulische doorlatendheid is de meest dominante en de meest heterogene parameter die de stroming van grondwater op veldschaal beïnvloed.

Laboratorium experimenten geven metingen van de parameters van subeenheden op een fijne schaal. Als laboratorium metingen stochastisch worden gebruikt in het geologische model dan moet het structurele model van de ondergrond op de zelfde schaal als die van de metingen worden opgesteld. Stromingsberekeningen met deze resolutie op veldschaal zijn onhanteerbaar en daarom is hier een nieuwe adaptive multiscale techniek ontwikkeld. Terwijl constitutieve relaties kunnen veranderen bij overgangen naar verschillende schalen, wordt Darcy's wet geldig verondersteld op de laboratoriumschaal en de veldschaal.

Tegenwoordig worden opschalingsmethoden toegepast die, zowel functioneel als structureel, informatie overdragen over de hiërarchie van schalen. Dit gebeurt van de fijne schaal naar de grove schaal en niet vice versa, met effectieve of equivalente parameters. De doorlatendheid is geen additieve parameter en het is dus niet mogelijk om een equivalente doorlatendheid op de grove schaal te berekenen als een eenvoudig gemiddelde van metingen op een fijne schaal. De equivalente doorlatendheid wordt vaak bepaald op basis van een stromingscriterium of een energiedissipatie criterium. De equivalente doorlatendheid komt alleen overeen met de effectieve doorlatendheid als de variatieschaal veel kleiner is dan de schaal waarop de doorlatendheid wordt bepaald. Deze effectieve doorlatendheid is een constante tweede orde tensor parameter. De equivalente doorlatendheid is echter niet uniek en is afhankelijk van de randvoorwaarden van het monster. Effectieve doorlatendheid geldt alleen voor discrete hiërarchische systemen waarvan de schalen kunnen worden gescheiden.

De nieuw ontwikkelde multischaal techniek breidt de oorspronkelijke multischaal eindige elementen methode op twee niveaus uit naar een hiërarchie van schalen. Met multischaal eindige elementen methoden wordt het fijne schaal gedrag door middel van multischaal basis functies vastgelegd op de grove schaal. De gewichten van de

multischaal basis functies volgen uit locale stromingsberekeningen. De procedure verwijdt alle fijne schaal knopen uit het subdomein waardoor echter fouten worden geïntroduceerd aan de randen van het subdomein. De op deze wijze geformuleerde methode vormt een klasse van subdomein decompositie technieken. Het kan worden aangetoond dat een sequentiële implementatie niet sneller is dan een optimale solver zoals de multigrid methode. De methode leent zich echter goed voor een parallelle implementatie.

De voorgestelde methode is gebaseerd op een conforme eindige elementen formulering voor simplex elementen. De conforme eindige elementen methode formuleert behoud van massa per knoop, continuïteit van flux over de element randen wordt niet behouden. Zones waarin grote fouten in de snelheden over de elementranden optreden worden met een mesh verfijningscriterium gedetecteerd, waarna een adaptieve verfijningsprocedure het mesh lokaal verrijkt zodat de fout wordt gecorrigeerd. Multiscale basisfuncties volgen uit de oplossing van locale stromingsproblemen over patchen van simplex elementen. Lineaire randvoorwaarden worden opgelegd op de randen van uitgebreide subdomeinen, waarmee randeffecten op de patch worden gereduceerd. Met deze procedure wordt het gedrag op de grove schaal nauwkeuriger bepaald. Uitgebreide locale stromingsproblemen introduceren echter discontinuïtetten in de basisfuncties en genereren nieuwe relaties tussen knopen op de grovere schaal. Daarom wordt de voorkeur gegeven aan het oplossen van het lokale probleem met randvoorwaarden die volgen uit een dimensioneel gereduceerd probleem. Een tweede verfijningscriterium vergelijkt functiewaarden op patches met waarden op uitgebreide subdomeinen. Druk-dissipatie middeling benadert de multiscale operator op het grove grid en ondersteunt een functionele adaptieve formulering. De multiscale middelingsmethode berekent equivalente doorlatendheidstensor componenten en reproduceert de ijle matrix op de grove schaal. De belastingsgevallen voor de locale problemen volgen uit het sommeren van multiscale basis functies. De multiscale basis functies extrapoleren de oplossing van de grove schaal naar de fijne schaal. Op deze schaal worden discontinuïtetten in het snelheidsveld gedetecteerd en vergeleken met het eerste criterium. De berekende equivalente doorlatendheid wordt gebruikt in het raamwerk van een geometrische multigrid methode bij het berekenen van grove schaal operatoren. Met de hiërarchie van multiscale basis functies, die de druk op ieder grof niveau relateert aan het volgende fijne niveau, worden de intergrid transfer operatoren gegenereerd.

De voorgestelde aanpak geeft een robuust en efficiënt algoritme, gebaseerd op de multiscale eindige elementen methode. Met het algoritme kan stroming door gedeeltelijk verzadigde hiërarchische heterogene formaties, en het daaraan volledig gekoppeld transport van opgeloste stoffen en warmte, worden gesimuleerd. Met de multiscale eindige elementen formulering worden numeriek gehomogeniseerde discrete stromingsvergelijkingen gegenereerd en worden lokaal verfijnde snelheidsvelden bepaald, die nauwkeurige transportberekeningen ondersteunen. De prestatie van het algoritme wordt met een aantal realistische cases geïllustreerd.

John van Esch.

# Curriculum vitae

John van Esch was born on the 10<sup>th</sup> of June 1966 in Heemstede.

In 1985 he started his Civil Engineering study at Hogeschool Haarlem. He was a practical trainee for a period of one year, firstly at water company 'Gemeentewaterleidingen Amsterdam' and secondly at geotechnical consultancy 'Tjaden Grondmechanica'. He obtained a B.Eng. degree in 1990. In 1989 he became a student at the faculty of Civil Engineering of Delft University of Technology. He specialized in geo-technology and hydrology, graduated (cum laude) in 1992 obtaining a MSc degree. During the following years, he participated in a number of advanced courses at Stanford University, Wageningen University, IHE Institute for Water Education, Wessex Institute of Technology, and Universitat Politècnica de Catalunya.

He started to work as a research engineer in the field of subsurface hydrology at Grondmechanica Delft, Department of Mathematics and Informatics in 1992. In this position, he collaborated in a research program on the assessment of time dependent stability behavior of embankments, and developed numerical tools for operating geo-hydraulic control systems, for analyzing time series from well measurements, and for analyzing retaining wall constructions.

In 1995, he became a research assistant at the Geotechnical Laboratory of Delft University of Technology where he carried out research on a part time basis, for a period of five years. During this period he focused on computational subsurface flow. At GeoDelft, he switched to the Department of Strategic Research. Next to geo-hydrological model applications and geo-statistical analysis, he participated in a EU-project on the interaction of subsurface hydrology and stability of landslides. He co-worked with Plaxis for about two year on the development of the unsaturated groundwater flow simulator Plaxflow.

In the period of 2005 till 2007, he worked at GeoDelft's Department of Soil structures. He participated in a research program on the development of a geo-hydrologic design procedure for peat dikes under drying and wetting conditions, and developed an algorithm for functional analyses of jetgrout bodies.

In 2008, he became a research engineer in the field of computational groundwater mechanics at the Deltares Software Center. He co-works with Universität Stuttgart on the development of a consolidation module, based on the material point method. In 2009 he switched to Deltares' foundation and subsurface structure unit and studied two-phase dynamics.