Soil Physics

Lecture Notes

v2.2β, Autumn 2012

Kurt Roth

The current version of these notes is available from

`www.iup.uni-heidelberg.de/institut/forschung/groups/ts/students/sp`

Please

Do not put a copy of these notes on your homepage
but rather link to the original site

and cite them as

Roth, K., 2012: *Soil Physics. Lecture Notes.*
Institute of Environmental Physics, Heidelberg University

giving the above link whenever possible.

Contact

Kurt Roth
Institute of Environmental Physics, INF 229
Heidelberg University
D-69120 Heidelberg

`kurt.roth@iup.uni-heidelberg.de`

Cover

Irrigated agricultural field in the North China Plain near Fengqiu, Henan Province, China.

Preface

Soil physics aims at understanding physical processes in terrestrial environments and at the quantitative prediction of the associated phenomena. This is challenging because the underlying processes are often nonlinear and coupled, and they run in hierarchical heterogeneous media with hardly known architectures. These media – porous structures with intricate geometries that originate from various generators operating on different length scales – set terrestrial systems apart from other large environmental compartments like the atmosphere or the ocean. In these other compartments, multi-scale phenomena almost exclusively arise from the dynamics of the process, typically from fluid dynamics, whereas the boundary conditions are irrelevant to a great extent.

While the physics of most processes relevant for soils is understood reasonably well at the scale of individual pores, this is no more the case at the larger scales of a soil profile. As scales increase even further, eventually to entire landscapes, ever more heuristic approximations have to be introduced in order to arrive at palpable results. These in turn are mandatory for understanding and handling a great many environmental and engineering problems.

These notes attempt to explore the transition from fundamental understanding to heuristic description and quantitative approaches to larger scales. The thematic focus is on transport processes, in particular on subsurface flow of water, on transport of dissolved substances, plant nutrients as well as groundwater contaminants, and on the movement of thermal energy. These processes are studied at scales ranging from a few hundred micrometers to a few hundreds of meters. Much smaller and much larger scales are only considered occasionally. Notable omissions in the spectrum of physical processes include soil mechanics, soil erosion, surface water, and soil-vegetation-atmosphere interactions. Also, experimental methods are not represented to the depth that would do justice to their importance in a still essentially explorative discipline. I should acknowledge and emphasize that transport processes, while being an essential and in some sense primary aspect of soil systems, are but one class of processes relevant for the function of soils in the environment. Much larger classes are formed by chemical interactions, microbiological processes, and by the actions of macroflora and -fauna.

The notes are written from a rather fundamental perspective but attempt to keep applications in mind. They try to strike a balance between clean mathematical formulations and simple intuitive concepts, with necessarily many compromises on both ends. The material is aimed at advanced undergraduate and graduate students in environmental sciences, physics, and engineering and it does require some passion for mathematics. Some problems are added at the end of each chapter. They fall into two classes: (i) simple calculations, just applying some equations given in the text, or marginal extensions of the theory and (ii) qualitative explorations of complicated phenomena which demand bold assumptions and may only yield order of magnitude answers. The first class is aimed at exercising the mechanics of the theory while the second one is to entice the reader to plough deeper.

Although the roots of these notes reach back to 1993, they continue to evolve, with the material still not adequately balanced and certainly containing a number of suboptimal explanations, mistakes, and errors, hopefully not too many. I am most grateful for any and all suggestions on how to correct and improve things such that they best serve the reader.

How to Work through these Notes The chapters following the introduction are arranged such that fundamental processes are studied first, bringing there representation from the pore scale to the continuum scale, and only then are they explored as they work in soils. However, the chapters are rather independent of each other, with ample cross-references where previous results are required, such that they may be worked through along different lines. With the priority on the continuous development of the thematic fields, the sequence 2, 3, [5], 6, 4, 7, 8 may be optimal, whereas with a priority on increasing difficulty, the sequence 3.1, 3.3, 4, 5, 7, [8], 2, 3.2, 6 is more appropriate.

Acknowledgements I thank a number of colleagues for providing critical comments, data sets, images, and numerical solvers: Hans-Jörg Vogel, now at Umweltforschungszentrum (UFZ), Halle, for the images on soil cross-sections, reconstructed three-dimensional representations, and innumerable discussions on soil hydraulics; Ute Wollschläger, now at UFZ, Leipzig, for many of her research results on groundwater flow in the Rhein-Neckar region; Dietmar Wagenbach, Institute of Environmental Physics (IUP), Heidelberg University, for critically reading the chapter on fluid dynamics; Ed Sudicky, University of Waterloo, for the Borden data set; Isabelle Cousin, Institut National de la Recherche Agronomique, Orléans, for the three-dimensional dataset from microscopic cross-sections of a soil sample. Whenever satellite images were needed for illustration, `modis.gsfc.nasa.gov` was my source. The numerical simulations of groundwater flow were run with an AMG (algebraic multi-grid) solver developed by Peter Bastian and Olaf Ippisch, Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University. Those of flow in the vadose zone were done with `SWMS`, the precursor of `HYDRUS`, by Jirka Šimůnek, UC Riverside.

Finally, I am glad to acknowledge my students for everything from pointing out errors to suggesting improvements. Most importantly though, I appreciate their interest, critical questions, and inspiring after-class discussions which makes teaching so rewarding.

Heidelberg, October 2012 Kurt Roth

Contents

List of Figures

List of Tables

Symbols

This list contains the most important symbols and notations used. It is still rather incomplete and will be upgraded in a future version. Where possible, a reference (equation and page number) is given and the unit is indicated in brackets.

The mathematical structure of symbols is indicated by their typographical appearance:

a	scalar
$\mathbf{a}$, $\widehat{\mathbf{a}}$	vector, unit length vector
A	tensor
$\sin$	standard function

Subscripts usually refer to a component of a vector (x, y, z, or 1, 2, 3) or to the phase (g, w, s). When dealing with quantities that refer to different scales, the superscripts μ and m are used to indicate the micro- and the macroscale, respectively. They are suppressed when only one scale is considered.

Parentheses $(\ldots)$ are used for arguments of functions and operators, brackets $[\ldots]$ group terms or indicate concentrations of a chemical component, and curly brackets $\{\ldots\}$ collect operators.

The arguments of functions are suppressed if they are clear from the context, e.g., $\partial_x f$ instead of $\partial_x f(x)$. They are written, however, if the dependence on an argument is emphasized, e.g., $K_d(C_w)$ for a nonlinear adsorption isotherm. Similarly, derivatives are written as shorthand operators: ∂_x, ∂_{xx} for the first and second partial derivative with respect to x, respectively, and d_t for the total derivative $\partial_t + \mathbf{v} \cdot \nabla$ with respect to time. If the derivative is to be emphasized, the long form $\frac{\partial \diamond}{\partial x}$ is used.

Sign Convention

The normal vector $\mathbf{n}$ on the surface of some volume points outwards. The vertical (z) axis points downward, in the direction of the acceleration of gravity. Its origin is typically chosen at the soil surface. Accordingly, z is called the depth.

Lowercase Latin Symbols

$\mathbf{g}$	acceleration of gravity $[\mathrm{m\,s^{-2}}]$		
h_m	matric head $[\mathrm{m}]$	3.41	62
$\mathbf{j}$	flux; quantity X flowing through unit area per unit time $[\mathrm{X\,m^{-2}s^{-1}}]$		
$\mathbf{j}_s$	mass flux of solute $[\mathrm{mol\,m^{-2}s^{-1}}]$		
$\mathbf{j}_w$	volume flux of water $[\mathrm{m\,s^{-1}}]$	3.24	52
$\widehat{\mathbf{n}}$	unit normal vector		
pe	microscopic Peclet number	4.15	81
$\mathbf{q}$	flow $[\mathrm{X\,s^{-1}}]$		
$\mathbf{r}$	position $[\mathrm{m}]$		
$\mathbf{v}$	velocity $[\mathrm{m\,s^{-1}}]$		
$\mathbf{x}$	position $[\mathrm{m}]$		

Uppercase Latin Symbols

$\mathbf{A}$	hydraulic gradient	5.9	105
Ca	capillary number	6.48	189
C_m	concentration in mobile phase $[\mathrm{mol\,m^3}]$	7.40	210
C_im	concentration in immobile phase $[\mathrm{mol\,m^3}]$	7.40	210
C_t	total concentration $[\mathrm{mol\,m^3}]$		
C_s	concentration in sorbed phase $[\mathrm{mol\,kg^3}]$		
C_w	concentration in water phase $[\mathrm{mol\,m^3}]$		
D	dispersion tensor $[\mathrm{m^2s^{-1}}]$	6.4	135
$\mathbf{F}$	force $[\mathrm{kg\,m\,s^{-2}}]$		
Fr	Froude number	2.24	22
K	hydraulic conductivity tensor; $[\mathrm{J^{-1}m^5s^{-1}}]$ in potential form or $[\mathrm{m\,s^{-1}}]$ in head form	3.32	55
k	permeability tensor $[\mathrm{m^2}]$	3.30	54
R	retardation factor	7.43	211
Re	Reynolds number	2.24	22
$\mathcal{S}$	linear system	7.34	207
Pe	Peclet number	7.25	203
St	Strouhal number	2.24	22
$\boldsymbol{\mathcal{V}}$	generalized velocity operator	4.42	90

Lowercase Greek Symbols

ψ_w	water potential, potential energy of water [J m^{-3}]	6.1	135
ψ_m	matric potential [J m^{-3}]	3.19	51
ρ_b	bulk density of porous medium [kg m^{-3}]	3.13	49
ρ_m	mass density of solid matrix [kg m^{-3}]	3.14	49
ρ_w	mass density of water [kg m^{-3}]	3.21	52
θ	volume fraction [–]	3.11	46
θ_m	mobile water content [–]	7.40	210
θ_im	immobile water content [–]	7.40	210
θ_w	volumetric water content [–]	3.23	52

Uppercase Greek Symbols

Θ	saturation [–]	3.15	49

Mathematical Notation

erf	error function
erfc	complementary error function
var	variance
∂_t	partial derivative with respect to time [s^{-1}]
∂_x	partial derivative with respect to spatial coordinate x [m^{-1}]
∂_{xx}	second partial derivative with respect to spatial coordinate x [m^{-2}]
d_t	total derivative with respect to time [s^{-1}]
∇	partial derivative with respect to space [m^{-1}]
∇_h	horizontal partial derivative with respect to space [m^{-1}]
$\widehat{f}$	Laplace transform of function f
$\mathcal{L}$	Laplace transform
$\mathcal{L}^{-1}$	inverse Laplace transform
$\diamond$	a generic argument

1

Introduction

Our *physical environment* may be divided roughly into three spheres: the atmosphere, the oceans, and the terrestrial environment. The latter may be further subdivide into (i) soils and sediments, (ii) snow, glaciers and ice shields, (iii) vegetation, and, depending on the perspective, (iv) surface waters like rivers and small lakes. Soils and their icy analogues are structurally similar in that they are both solid porous media whose pore space is filled with liquid water and with air. The main difference between them is that one is of predominantly mineral origin while the other consists almost exclusively of frozen water. They both may be considered as a boundary region between the fluid compartments – atmosphere and oceans – and the solid lithosphere, possibly a rather thick one. Vegetation may also be looked upon as another boundary, one between soil and atmosphere. While its most important function is arguably the assimilation of solar energy into carbon compounds by consuming CO_2 and producing O_2, its next important function is the increase of the coupling between soil and atmosphere. This is accomplished on the one hand by the rooting system which pervades the top soil layers and on the other by the stems and foliage that tap into the atmosphere. The extent of the soil-atmosphere boundary layer, which is measured in centimeters for bare soil, thereby extends to meters or even tens of meters with vegetation. This in turn leads to a dramatic increase of the water, carbon, energy, and momentum fluxes between soil and atmosphere. Finally, the inclusion of small surface waters into terrestrial systems is debatable since they clearly share important properties with oceans. However, they are also akin to structures found in some sedimentary units, in particular in karstic regions with their extended cave systems with openings ranging from a few millimeters to hundreds of meters.

In these notes, after outlining the general global context, we focus on soils and there primarily on the flow of water, the transport of dissolved chemicals, and the movement of heat. Important aspects which will not be covered here include soil mechanics and the dynamics of soil structures, erosion and sedimentation, and the very genesis and evolution of soils. Soils provide a number of crucial ecosystem functions. To name a few, they store water for the eventual uptake by plant roots, moderate local temperature through the latent heat of evaporating water, and act as highly efficient filters and reactors which lead to clean groundwater. The importance of soils is for instance illustrated by *Churkina and Running* [1998] who found that net primary production – the net growth of plants – is limited by the availability of water for about 50% of the land surface and by low temperatures for about 30%. Obviously, there is some overlap between these, most importantly on the Tibetan plateau which is dry *and* cold over large stretches. Still, soils are a critical factor in net primary production, hence in the fluxes of water, energy, and carbon between soil and atmosphere, for about 2/3 of the Earth's land surface.

1.1
Global Terrestrial Water and Energy Fluxes

Water, as a liquid component, is almost always on the move, typically in intricate and entangled cycles. We first look at terrestrial water fluxes at the global scale (Figure 1.1). The *global water cycle* transports water from the oceans through the atmosphere over the continents where it precipitates and eventually returns to the ocean as continental runoff. Exclude Antarctica, the average flux is 348 $\mathrm{mm\,y^{-1}}$ which corresponds to a flow of $47\cdot10^3$ $\mathrm{km^3y^{-1}}$ distributed over the continental area of $134\cdot10^6$ $\mathrm{km^2}$.

In the *internal land-atmosphere water cycle* 58% of the total precipitation of 836 $\mathrm{mm\,y^{-1}}$ returns to the atmosphere via *evapotranspiration* (ET). This flux may be decomposed into three major components: *evaporation* from the soil surface (176 $\mathrm{mm\,y^{-1}}$), *transpiration* of soil moisture taken up by plant roots and transported to the leafs where it evaporates (233 $\mathrm{mm\,y^{-1}}$), and *interception*, precipitation that did not reach the soil surface but directly evaporates from the leaf surfaces (80 $\mathrm{mm\,y^{-1}}$).

We notice in passing that sometimes "blue" and "green" water is distinguished, particularly for the purpose of water resources management. Blue water refers to rivers and lakes while green water refers to that in the subsurface which can, potentially, be directly tapped by the vegetation [*Falkenmark and Rockström* 2006].

Earth's water cycles are driven by solar radiation (Figure 1.2). The total incoming short-wave energy flux received from the sun outside of the atmosphere is some 342 $\mathrm{W\,m^{-2}}$. Over the land surface, again excluding Antarctica, some 60% of this flux is on average scattered back into space either by the atmosphere – by clouds, aerosols, and molecules – or by the land surface (albedo). Eventually, 140 $\mathrm{W\,m^{-2}}$ is absorbed by the land surface.

The incoming short-wave radiation is counterbalanced by (i) outgoing long-wave (thermal) radiation with a net flux of 66 $\mathrm{W\,m^{-2}}$ and (ii) convective heat fluxes totaling 72 $\mathrm{W\,m^{-2}}$. The *net thermal radiation* is actually the difference of two rather large fluxes: the thermal emission of the Earth surface with some

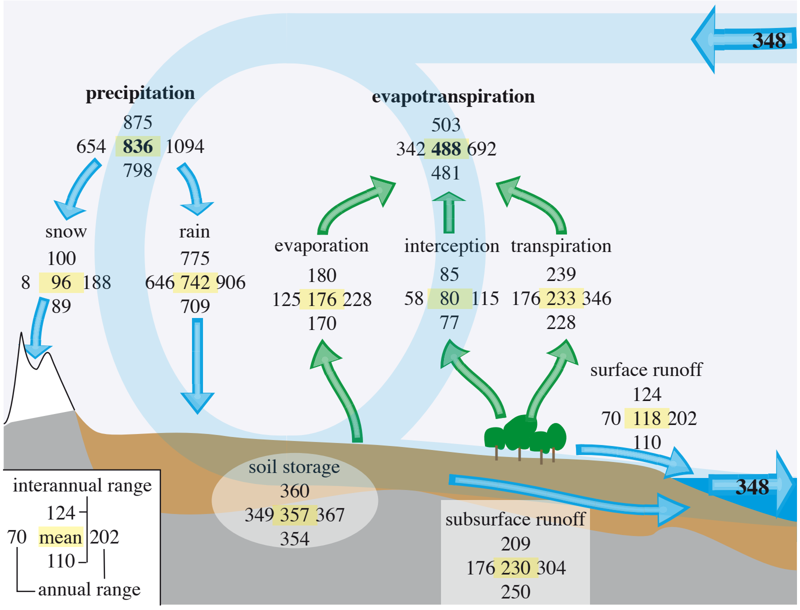


Figure 1.1. Schematic of global terrestrial water cycle [after *Dirmeyer et al.* 2006]. The numbers have been calculated from simulations of the 10-year period 1986-1995 by 15 different global circulation models. In addition to the average numbers for the entire period (yellow rectangles), the range of fluctuations within and between years is indicated. Fluxes are in mm y^{-1}, soil storage is in mm. The numbers represent averages over the continental area, excluding Antarctica. With an area of $134 \cdot 10^6$ km^2, a flux of 1 mm y^{-1} corresponds to a flow of 134 $km^3 y^{-1}$.

390 W m^{-2}, of which a small fraction, some 40 W m^{-2}, is emitted directly into space through the "atmospheric radiative window", and back radiation from greenhouse gases (H_2O, CO_2, CH_4,. . .), aerosols, and clouds [*Kiehl and Trenberth* 1997]. Also the *convective heat flux* may be separated into two components, the *sensible heat flux* with some 33 W m^{-2} in the form of raising warm air and the *latent heat flux* with some 39 W m^{-2} which represents the energy stored in the evaporated water that is released upon condensation higher up in the atmosphere. This last energy flux drives the water cycle through evaporation, transport, and condensation of water. The fluxes of latent heat, 39 W m^{-2}, and of evaporated water, 488 mm y^{-1} (Figure 1.1), are linked through the latent heat of evaporation, 2.5 MJ kg^{-1}.

Two small fluxes remain to be mentioned. The ground heat flux of 1 W m^{-2} represents global warming during the 10-year period 1986-1995 considered. It leads to a net heat flux into the ground, opposite to the geothermal flux of some 0.06 W m^{-2}. The "other fluxes", some 0.8 W m^{-2}, originate for instance from the melting of precipitated snow.

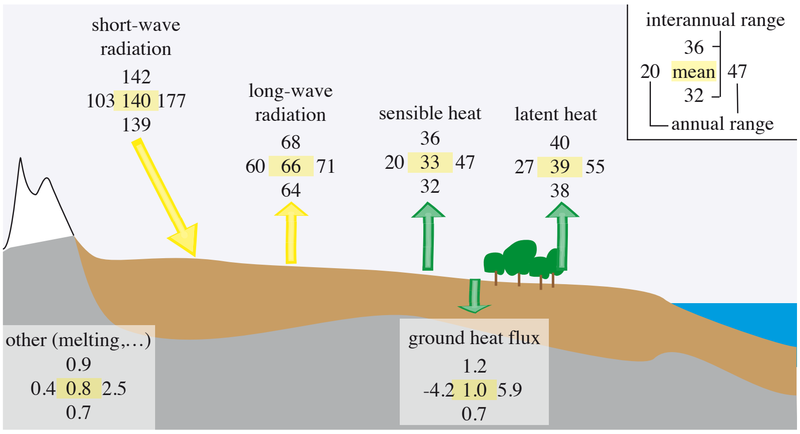


Figure 1.2. Schematic of global terrestrial energy fluxes [$\mathrm{W\,m^{-2}}$] after *Dirmeyer et al.* [2006]. The numbers have been obtained from the model runs that already led to Figure 1.1 and are presented in analogy.

A caveat has to be added to the numbers given in Figures 1.1–1.2. The corresponding fluxes can neither be measured directly nor calculated in a simple manner from other measurements, hence are not known objectively. Instead, they have been obtained as ensemble averages from 15 different global circulation models (GCMs) which (i) represent our best current understanding of the complicated processes in the terrestrial environment, the atmosphere, and the ocean and which (ii) have been calibrated with available measurements, hence include our best current quantitative knowledge.

We finally notice that such cycles and average fluxes are gross simplifications of an exceedingly complicated reality. Local fluxes may easily deviate by an order of magnitude from the global average – just think of the Sahara, the Kongo basin, and Siberia – and the cycles in Figure 1.1 disintegrate into a multi-scale hierarchy of entangled flows as we move to higher resolutions of reality.

1.2
Anthropogenic Use of Water

Our society uses and consumes water in many different ways. A very small fraction is consumed directly for drinking and cooking but the largest fraction goes into crop production on irrigated land. Further usage includes a wide spectrum of industrial productions and municipal waste removal. In the following, we will only consider water going into agricultural production since this is the dominating and still increasing fraction of anthropogenic water use, it is the major limiting factor in food production, and it is a strong motivation to study water movement in the terrestrial environment.

1.2.1
Water Quantity

Globally, some 40% of all crop is produced on irrigated land with increasing tendency. These crops are primarily used for food, both for people and for animals. An increasing fraction is transformed into biofuel, however. For instance, by 2005, one sixth of the US corn production was converted to ethanol which then supplied some 2% of the domestic demand for fuel. In Brazil, even one half of the sugarcane production was processed to ethanol. One effect of this growing demand are increasing food prices: the price of corn in the US doubled within a year, for instance. Another one is the unabated pressure on water resources for irrigation.

Estimates for the water consumption of some goods are given in Table 1.1. Worldwide, some $2.4 \cdot 10^6$ km^2 of land are irrigated with an average of 1.2 $\mathrm{m\,y^{-1}}$. This leads to a water demand for irrigation of some $2.9 \cdot 10^3$ $\mathrm{km^3 y^{-1}}$, of which 50…80% is lost to evapotranspiration [*Postel et al.* 1996]. Translated into the units of Figure 1.1, demand for irrigation corresponds to 22 $\mathrm{mm\,y^{-1}}$ or 2.6% of the total precipitation on the continents except Antarctica. The picture becomes more alarming, however, when we realize that only a small fraction of the total precipitation is available for anthropogenic use. This is essentially the average recharge of aquifers and a fraction of the river flow in inhabited areas. Storm-flow, for instance, is only usable when large intermediate storage is available, mostly dammed valleys, and flow through uninhabited land would have to be redirected for usage. As an aside, we notice that in many regions, groundwater is pumped at a rate that exceeds recharge which is not a sustainable practice. This so-called water mining is not considered in the following. *Postel et al.* [1996] estimated

Table 1.1. Total consumption of water – kg of water per kg of product, often called "virtual water" – for the production of various crops and meat (adapted from *United Nations* [2006]). The numbers are global estimates deduced from statistical information for individual countries. Depending on crop, irrigation technology, and climate, large deviations from the average value occur. Rice, for instance requires between 900 kg/kg in Belgium and 15'000 kg/kg in Sudan [*Hoekstra and Hung* 2002].

product [kg]	water consumption [kg]
potatoes	160
maize	450
wheat	1'200
rice	2'700
poultry	2'800
pork	6'000
beef	16'000

Figure 1.3. Huang He (Yellow river) at Lanzhou, with some 2000 km to go to the sea, in September 2004. This is the second longest river in China and number 6 in the world. During the decade 1990-1999, this stream did not reach the sea for extended periods of 50 to 200 days but trickled out between 200 and more than 600 km inland [*Fu et al.* 2004].

the presently available runoff as $12.5 \cdot 10^3$ km^{3}y^{-1}. Hence a realistic estimate for the water demand for irrigation is some 23% of the sustainably available freshwater.

The impact of irrigation on natural waters strongly varies between regions and seasons. This is exemplified by the Huang He, the Yellow river, in northern China (Figure 1.3) with regularly trickles out far inland. A large multi-decadal water project, the "*South-North Water Transfer Project*", is currently under way to relieve the situation by diverting water from the Chang Jiang (Yangtse) catchment, Chinas largest river, to that of the Huang He. The Colorado river in the USA and Mexico, the Nile river in North Africa, or the Amu- and Syr-Darya in the Aral region (Figure 1.4) have a very similar fate: all of them are consumed to a large degree by irrigation projects and the remaining flow is typically heavily polluted with agricultural and industrial contaminants.

Tapping into an essential natural cycle to an extent as is done with the water cycle raises the question of responsible management, and in particular the question of the most efficient use of water. One phenomenon that must be understood and predicted is the falling of water tables which result in desertification, when natural vegetation does not reach the water anymore, or to salt-water intrusion in coastal regions. Another, equally dangerous phenomenon are raising water tables due to large amounts of imported irrigation water. This may lead to swamping and, particularly in dry regions, to salinization which may irreversibly destroy the soil's capacity to carry vegetation. Often, regional climatic changes are associated with both of these phenomena as is for instance the case in the Aral region where a moderately humid and partly forested region turned into a desert with frequent dust storms sweeping the region. In contrast, the Central Valley of California became more humid through the import of large amounts of water from other catchments. Clearly,

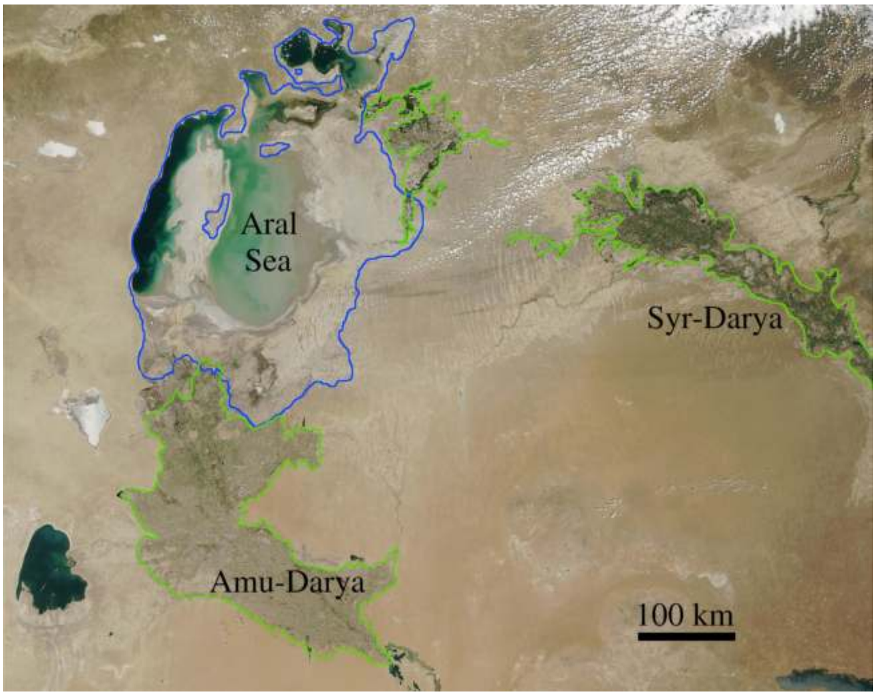


Figure 1.4. Aral region as seen from the MODIS sensor aboard Terra on December 2002. The blue line indicates the perimeter of the Aral Sea in 1970 and the green lines outline irrigated land. *Satellite image: MODIS Land Rapid Response Team, NASA/GSFC*

attempts to alleviate adverse effects of older irrigation projects and to prevent them for new projects demand a quantitative understanding of the flow of water through the subsurface.

1.2.2
Water Quality

So far, we focussed on the quantity of available water. This issue cannot be regarded without simultaneously addressing quality, however, since a minimal quality of the water is required for most uses. In the first place, this concerns salinity, the total salt load, irrespective of the salt's composition. For instance, most crops can tolerate a salinity of $1\ \mathrm{g}\,\ell^{-1}$ but only very few can cope with $6\ \mathrm{g}\,\ell^{-1}$. As a reference, the mean salinity of the oceans is $34\ \mathrm{g}\,\ell^{-1}$. Sea water is thus utterly useless for irrigation. The quality requirements become much more strict for specific toxic substances, in particular when the water is potentially used for direct consumption like our tap water. Examples are the World Health Organization's (WHO) limit on arsenic which is $10\ \mu\mathrm{g}\,\ell^{-1}$ or the German limit for the concentration of the sum of all

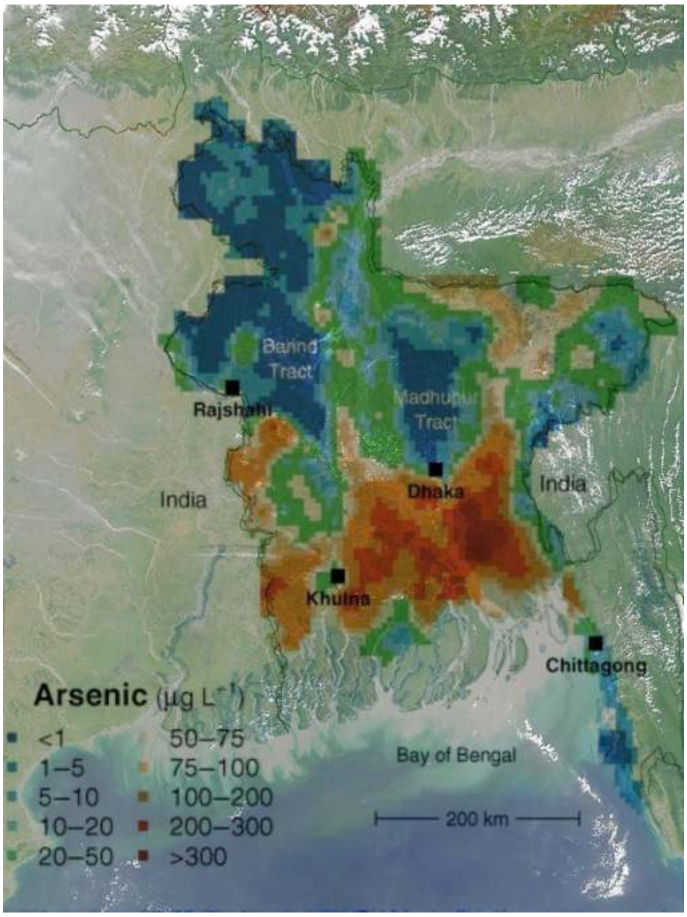


Figure 1.5. Arsenic distribution in the groundwater of Bangladesh drawn as an overlay of the concentration estimates from *Smedley and Kinniburgh* [2002] over a satellite image of the region as seen from the MODIS sensor aboard Terra in October 2002. The World Health Organization's limit on arsenic in drinking water is 10 $\mu g\,\ell^{-1}$. *Satellite image: MODIS Land Rapid Response Team, NASA/GSFC*

pesticides in drinking water which is 0.5 $\mu g\,\ell^{-1}$. Quality requirements may thus severely limit the available quantity of water.

Water is an almost universal solvent and thus contains a huge diversity of dissolved substances, particularly in developed regions. Quantitatively, the major sources of solutes are (i) rocks and soil minerals, which are dissolved by the percolating water, and (ii) the ocean primarily through spray. Such completely natural processes may actually lead to quite severe environmental problems as is exemplified by the arsenic contamination of a large fraction of the aquifers in Bangladesh (Figure 1.5 and *Burgess et al.* [2010]). Human activity provides another important source of dissolved substances. The largest quantity again comes from agricultural production which releases huge amounts of fertilizers, pesticides, and in recent years even hormones into the environment. A fraction of these substances eventually ends up in surface- and groundwater. Then there is the plethora of substances released by industrial and individual activities: organic compounds like the gasoline additive MTBE, gasoline itself and other hydrocarbons from spills, as well as inorganic substances like heavy metals from mine tailings, abrasion of tires, and waste dumps or the emissions of sulfur-, nitrogen- and carbon-oxides into the air which are eventually dissolved in precipitation.

Obviously, there also exist efficient cleaning mechanisms which may be roughly grouped into (i) phase changes and (ii) reaction-filtration processes. Phase change is dominated by evaporation from ocean and land surfaces with the formation of sea ice in polar regions being of secondary importance. Cleaning is accomplished by excluding solutes, hence by concentrating them in the liquid phase. As concentration increases, some solutes may actually precipitate, form a solid phase, and thus be removed from the water cycle. Reaction-filtration acts most intensely in soils, and in particular in the top layer with a thickness of $0.1 \ldots 1$ m. Here, the porous matrix provides an effective filter for large particles like bacteria, viruses, and, to a lesser degree, colloids. Smaller particles like ions may be retained by the electrical surface charges that are typical for the soil matrix. Most soils in temperate climates exhibit a negative surface charge that originates from clay minerals. They strongly adsorb cations like NH_4^+, PO_4^+, K^+, and a range of heavy metals like lead, cadmium, or platinum. While this cleans percolating water, it does not permanently remove anything. Indeed, the load of filtered and adsorbed substances may be released by erosion or by changing chemical conditions, in particular if pH-values decrease because of acid rain and thereby cause surface charges to change their sign. Besides filtration and adsorption, the biologically active top layers of soils form a very efficient reactor for the decomposition of most organic compounds, either chemically or, more often, microbiologically. The rich spectrum of microorganisms in soils and their adaptability to changing environmental conditions and substances is a major cleaning agent in the water cycle. As an example, consider pesticide application in Germany. In 1995, a total of some $35 \cdot 10^6$ kg was applied to an area of some $2 \cdot 10^5$ km^2. With an average seepage of $0.25\ \mathrm{m\,y^{-1}}$ – volume of water seeping into groundwater per unit area of soil surface – and presuming negligible sorption and decay, we calculate an average pesticide concentration in the groundwater of some $0.7\ \mathrm{mg}\,\ell^{-1}$. Some of this pesticide mass is indeed found in the ground- and surface waters. A country-wide study between 1990 and 1995 estimated that some 30% of the groundwater was contaminated by pesticides or by their decay products and found some 10% of the measured concentrations were larger than $0.1\ \mu\mathrm{g}\,\ell^{-1}$, the German legal limit for any single pesticide. While these results are alarming – 70% of all drinking water in Germany stems from groundwater – they also demonstrate that sorption and decay reduce pesticide concentrations on average by a factor of at least 10^3. After all, 90% of the measurements showed values below 0.1 μg ℓ^{-1}, corresponding to a reduction by a factor of more than $7 \cdot 10^3$, and some 70% of the measurements showed no contamination at all. The soil environment thus appears as a rather efficient reactor, at least on average and for those organic substances which are amenable to microbial decomposition. Apparently, this system may also fail, often because of insufficient contact time when water flow is high and contaminants pass the active zone in a short time. A notorious problem in this context is the so called *preferential flow* which occurs (i) along macroscopic voids like cracks, worm holes, and root channels, (ii) with flow fingers that originate in unstable infiltration fronts, or (iii) due to strong heterogeneities which lead to a rather thin network of flow channels. A sufficiently detailed and accurate description of the flow field is thus a prerequisite for understanding the movement of dissolved substances from the soil surface to groundwater and further on to some pumping well or other exit surface.

We mention in passing that microbial decay also occurs in deeper soil layers and in aquifers, although typically at much lower rates than in the top layers. This has lead to the concept of *natural attenuation* which presumes that, given enough time and transport distance, a large class of contaminants will eventually be decomposed by the natural microbial population, possibly after some initial evolution. An engineered variant, *enhanced natural attenuation*, attempts to provide the microbial population with an energy source, e.g., nitrate, to facilitate co-metabolic decomposition of the contaminants. Since such a feeding can only be done through adding the required substances to the water that will eventually flow past the population of interest, the flow field needs to be known with sufficient accuracy. While this is already difficult, the microbial population adds a further complication in that it may change the pore space, hence the hydraulic properties of the material, by its very growth. This nicely illustrates the challenges faced by quantitative models of subsurface transport. They must yield accurate predictions at a typically large scale of interest, from a few meters to man kilometers, while taking into account processes which operate at a very much smaller scale, here the clogging of the pore space with operates in the reaction zone with a thickness ranging from a fraction of a millimeter to a few millimeters.

1.3
Why Study Small-Scale Processes in Soils?

After looking into the large-scale environmental issues soils are involved in, the question arises: "Why should we study transport processes at very small scales instead of directly working at the scale of eventual interest?" The answer to this has many facets.

First and foremost, soils are thin with thicknesses on the order of a meter and the strongest gradients are orthogonal to the surface. Hence, while soils are a global phenomenon and contribute to many aspects of the global environment, the pertinent processes are rather local. Even if we extend "soils" to include sediments, and with them groundwater where flow is predominantly horizontal, most phenomena still occur at rather small scales on the order of a few kilometers. This is in stark contrast to the atmosphere and, to a lesser extent, to the oceans both of which can only be understood at the global scale.

To be sure, there exist descriptions of soils at scales that are very much larger than a few meters, actually all the way up to the global scale. Such descriptions are for instance required for our daily weather predictions or for the representation of the carbon cycle in climate models. They are highly heuristic, however, and little more than a regression summary of past experiences, very often even a fudge component that is adjusted such that the models fit best with observations. While this may suffice for predicting an essentially stationary system, it is not satisfactory for dealing with a system that may evolve into states we have never observed quantitatively before. Unfortunately, we have to accept that currently there is little to immediately ameliorate the situation: Upscaling, the transfer of small-scale understanding to larger scales, is still a thorny area of active research. However, ever more powerful computers may eventually open an avenue to directly incorporate our solid small-scale concepts into simulations of phenomena at the very much larger scales of interest.

It is furthermore worth mentioning that the processes illuminated in the following chapters, and the challenges encountered thereby, are highly relevant in a number of other fields. Examples include printing, where the infiltration of dyes into the porous texture of paper or cloth is of interest, or fuel cells, where the transport of gases into a porous catalytic structure and the removal of the resulting fluid is of crucial importance.

Finally, as a more fundamental comment, we appear to understand nature primarily at the microscopic scale. For the continuum, where the individual atoms are not visible anymore, this understanding is often cast into differential equations. Solving them then provides the vehicle to move up to larger scales. This works fine, and is a most powerful approach, as long as the phenomena at the larger scales are determined exclusively by the processes represented by the differential equations at hand. A prime example for this are the large fluid compartments, the atmosphere and the oceans. Natural porous media are an altogether different matter, however. As we will see, at each particular scale new structural elements typically enter the picture and modify the transport processes. Still, since we only understand these processes at very small scales, we have to start there and try to work our way up.

2
Fluid Dynamics

Fluid dynamics covers an often unsuspected range of phenomena that include the tea in your cup as well as a rushing river, molten magma, the ocean and the atmosphere, all the way to clusters of galaxies... and it also describes the flow of water through porous media. All these phenomena share a common structure in that they comprise a very large ensemble of "particles" – which may be anything from atoms to stars – that only interact locally during collisions which conserve particles and momentum, and which are isotropic [*Goldenfeld and Kadanoff* 1999]. The term "collision" is used in a weak sense here and merely indicates that, in contrast to a solid, particles are not bound to their position relative to their neighbors but are free to move and thereby interact with various other particles. This freedom to move is the defining characteristic of a fluid: it cannot maintain shear forces in static equilibrium.

For our purpose, we start at the scale where a fluid can be defined as a continuum object without the necessity to explicitly account for its atomistic nature. For slow motion in a large domain, the macroscopic state of such a fluid may be described by a few parameters the number of which equals the number of conserved quantities. For a simple fluid, conserved quantities are mass, energy, and momentum. Pressure p, temperature T, and velocity $\mathbf{v}$ are typically chosen as easily observable parameters. Their values at location $\mathbf{x}$ are defined from appropriate averages of the (large) ensemble of particles in the neighborhood Ω of $\mathbf{x}$. Prerequisites for this are that (i) Ω is much larger than the mean distance between particles, some 3 nm for air at standard temperature and pressure and about an order of magnitude smaller for water, (ii) Ω is much smaller than the spatial resolution of the macroscopic scale, and (iii) the time between collisions is much smaller than the temporal resolution of the macroscopic scale. With these premises it is useful to introduce the notion of a *fluid element*, a macroscopically infinitesimal volume of constant mass that is moved and deformed by the macroscopic velocity field $\mathbf{v}$. Notice that a fluid element is but an accounting volume and that its nature is rather ephemeral as its boundary becomes fuzzy due to molecular diffusion and its overall shape may be deformed strongly, e.g., in shear fields. For sufficiently slow flow fields, this does not cause difficulties, since the fluid element is only required to exist for an infinitesimal time, macroscopically speaking.

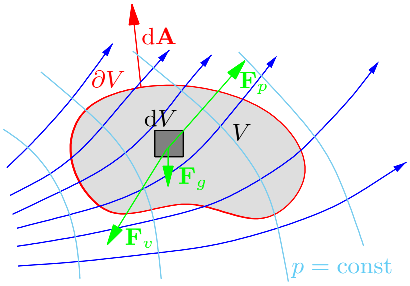


Figure 2.1. Fluid element with volume V, boundary ∂V and surface element $\mathrm{d}\mathbf{A}$. The forces acting on an infinitesimal sub-volume $\mathrm{d}V$ of the fluid element are shown in green. The pressure field which is the dominating driving force here is indicated by the cyan contours and the blue lines show some trajectories.

Modeling the dynamics of a fluid consists of three steps: (i) postulate forces and phenomenological equations that relate these forces to the resulting fluxes, (ii) formulate appropriate conservation laws, and (iii) specify appropriate thermodynamic relations – material properties – like the dependence of mass density on pressure and temperature, $\rho(p, T)$. Once the dynamics is formulated, typically as a set of partial differential equations together with some functional relations, initial and boundary conditions have to be specified such that they describe the initial state and the external forcing of the fluid, respectively. We will work through this program in the following.

2.1

Generic Formulation of Single Phase Dynamics

Consider the dynamics of an unbounded, isothermal, and isotropic fluid which is only subject to mechanical forces of gravity, pressure gradient, and friction.

2.1.1

Forces and Fluxes

Let V be an arbitrary fluid element with boundary ∂V and area element $\mathrm{d}\mathbf{A} = \widehat{\mathbf{n}}\mathrm{d}A$, where $\widehat{\mathbf{n}}$ is the outward pointing normal vector with unit length (Figure 2.1).

Gravity The force resulting from the acceleration of gravity $\mathbf{g}$ acting on the fluid element is

$$\mathbf{F}_g = \int_V \mathbf{g}\, \underbrace{\rho\, \mathrm{d}V}_{\mathrm{d}m} \; . \tag{2.1}$$

Pressure The hydrostatic pressure field p exerts the force

$$\mathbf{F}_p = -\int_{\partial V} p\, \mathrm{d}\mathbf{A} = -\int_V \nabla p\, \mathrm{d}V \; , \tag{2.2}$$

where the sign reflects $\mathrm{d}\mathbf{A}$ pointing outwards and Gauss' theorem has been used for the second equality.

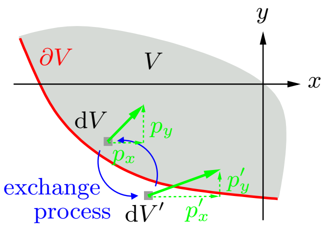


Figure 2.2.
Momentum flux across boundary ∂V of fluid element V. The infinitesimal volumes $\mathrm{d}V$ and $\mathrm{d}V'$ are exchanged by some microscopic process and carry their respective momenta $\mathbf{p}$ and $\mathbf{p}'$ with them. Hence, x-momentum is exchanged in x- *and* y- direction and similarly also y-momentum. This gives rise to the tensor $\mathsf{\Pi}$ of momentum flux. Exchange processes range from molecular diffusion to turbulence.

Friction Viscous forces develop in a fluid whenever velocity gradients exist. At the molecular level, viscosity results from the transfer of momentum when molecules are exchanged between high- and low-velocity regions due to thermal motion. As a generalization at the macroscopic level, we consider the exchange of infinitesimal volumes which leads to the *momentum flux* $\mathsf{\Pi}$ (Figure 2.2). A component Π_{ij} of the second rank tensor $\mathsf{\Pi}$ gives the amount of i-momentum that flows in j-direction through a unit area per unit time. A specific form of this tensor will be discussed in Section 2.2 below.

To calculate the viscous force, we recall that a rate of change of momentum corresponds to a force, i.e., $\mathbf{F} = \mathrm{d}_t\mathbf{p}$. Integrating the momentum flux Π_{ij} over the boundary ∂V of the fluid element, yields the total inflow of momentum per unit time, hence the rate of change of the fluid element's momentum. Thus, the viscous force that acts on the element becomes

$$\mathbf{F}_v = -\int_{\partial V} \mathsf{\Pi} \cdot \mathrm{d}\mathbf{A} = -\int_V \nabla \cdot \mathsf{\Pi}\, \mathrm{d}V \ , \tag{2.3}$$

where the sign again stems from $\mathrm{d}\mathbf{A}$ pointing outwards and the Gauss theorem was used for the second equality. In Cartesian coordinates $\{x, y, z\}$, the x-component of $\mathbf{F}_v$ is for instance given by

$$\begin{aligned} F_{vx} &= -\int_{\partial V} \{\Pi_{xx}, \Pi_{xy}, \Pi_{xz}\} \cdot \mathrm{d}\mathbf{A} = -\int_V \nabla \cdot \{\Pi_{xx}, \Pi_{xy}, \Pi_{xz}\}\, \mathrm{d}V \\ &= -\int_V \partial_x\Pi_{xx} + \partial_y\Pi_{xy} + \partial_z\Pi_{xz}\, \mathrm{d}V \ . \end{aligned} \tag{2.4}$$

2.1.2
Conservation Laws

For the situation we consider here, the conserved quantities of interest are mass and linear momentum.

Mass Conservation of fluid mass is conveniently formulated in the Eulerian framework where a volume element is considered that is fixed in space. With mass conserved, the rate of change of mass in volume V equals the mass flow across its boundary ∂V,

$$\underbrace{\partial_t \int_V \rho \, \mathrm{d}V}_{\text{mass in } V} = - \underbrace{\int_{\partial V} [\rho \mathbf{v}] \cdot \mathrm{d}\mathbf{A}}_{\text{mass flow across } \partial V} \overset{\text{Gauss}}{=} - \int_V \nabla \cdot [\rho \mathbf{v}] \, \mathrm{d}V \ , \tag{2.5}$$

where the sign stems from $\mathrm{d}\mathbf{A}$ pointing outwards. Since V is arbitrary, this implies

$$\partial_t \rho + \nabla \cdot [\rho \mathbf{v}] = 0 \ . \tag{2.6}$$

Momentum Conservation of linear momentum is most easily formulated in the Lagrangian framework, i.e., by moving with the fluid element. The momentum $m\mathbf{v}$ of the fluid element is $[\int_V \rho \, \mathrm{d}V]\mathbf{v}$, where $\mathbf{v}$ is the velocity of its center of gravity. Since by definition the mass of a fluid element remains constant, the rate of change of the momentum is $m\mathrm{d}_t\mathbf{v}$. This rate equals the sum of the external forces on the element, hence

$$\left[\int_V \rho \, \mathrm{d}V\right] \mathrm{d}_t \mathbf{v} = \mathbf{F}_g + \mathbf{F}_p + \mathbf{F}_v = \int_V [\rho \mathbf{g} - \nabla p - \nabla \cdot \mathsf{\Pi}] \, \mathrm{d}V \tag{2.7}$$

and further, since V is again arbitrary,

$$\rho \mathrm{d}_t \mathbf{v} = \rho \mathbf{g} - \nabla p - \nabla \cdot \mathsf{\Pi} \ . \tag{2.8}$$

This is the *Navier-Stokes* equation. For fluid flow through porous media, this is a rather general formulation. Notice however that fluid flow in other environments may demand modifications or additional terms. For instance, flow in the large fluid compartments of the Earth, the atmosphere and the oceans, is described by replacing $\rho\mathbf{g}$ by $-\rho\nabla\Phi$, where $\Phi(\mathbf{x}) = \int_\infty^{\mathbf{x}} \mathbf{g}(\boldsymbol{\xi}) \cdot \mathrm{d}\boldsymbol{\xi}$ is the geopotential, and the Coriolis force $-2\rho\boldsymbol{\omega} \times \mathbf{v}$ is added to account for the Earth's rotation with angular velocity $\boldsymbol{\omega}$. In a similar manner, electromagnetic forces may be included which leads to the field of magnetohydrodynamics that describes conservation of momentum in a plasma.

2.1.3
Dynamics

Combining (2.6) and (2.8) yields a formal description for the dynamics of a general fluid. Obviously, it is incomplete since there are, in three dimensions, 11 unknowns – ρ, p, the components of $\mathbf{v}$, and the independent components of the symmetric tensor $\mathsf{\Pi}$ – but only 4 equations. The remaining equations are supplemented by descriptions of material properties, i.e., by relating ρ and $\mathsf{\Pi}$ to p and $\mathbf{v}$.

2.2
Material Properties

At the beginning of this chapter, we chose to represent a fluid at the scale of a continuum, hence to disregard that in reality it is composed of atoms. This is successful because experience taught us that for most fluid flow phenomena it is quite sufficient to represent their physics at the scale of atoms in an average way. This typically leads to relations between continuum-scale variables, for instance between the fluid's mass density, temperature, and pressure. At least in principle, these relations could be calculated "from first principles", specifically from the interaction potential and the mass of the constituent atoms or molecules. However, since already small quantities of any substance involve a huge number of atoms, such calculations are not practical and it is more efficient to measure the desired relations and to parametrize the data.

In a single fluid system, pertinent material properties of the fluid are its density and viscosity which in general depend on temperature, pressure and, for liquids, solute concentration. In the following, we neglect the influence of solutes.

2.2.1
Density

We consider a fluid element and write its density as $\rho(T, p)$ with differential

$$\mathrm{d}\rho = \left.\frac{\partial \rho}{\partial T}\right|_p \mathrm{d}T + \left.\frac{\partial \rho}{\partial p}\right|_T \mathrm{d}p \ . \tag{2.9}$$

Introducing the coefficient of thermal expansion and the isothermal compressibility,

$$\kappa_T := -\frac{1}{\rho}\left.\frac{\partial \rho}{\partial T}\right|_p \quad \text{and} \quad \kappa_p := \frac{1}{\rho}\left.\frac{\partial \rho}{\partial p}\right|_T \ , \tag{2.10}$$

respectively, (2.9) may be written as $\rho^{-1}\mathrm{d}\rho = -\kappa_T \mathrm{d}T + \kappa_p \mathrm{d}p$. We notice in passing that the usual definition of the coefficients (2.10) refers to volume changes, e.g., $\kappa_p = -V^{-1}\partial_p V$. However, writing $\rho(T, p) = m/V(T, p)$, where m is the constant mass of the fluid element and V its volume shows that the two definitions are identical.

The compressibility of liquids is practically independent of pressure in the range of interest for fluid flow near the Earth surface. For constant temperature $T_0 = \text{const}$, we then obtain $\rho(T_0, p) = \rho(T_0, p_0)\exp(\kappa_p[p - p_0])$. For pure water at $T = 0°\mathrm{C}$, $\kappa_T \approx 10^{-4}\ \mathrm{K}^{-1}$ and $\kappa_p \approx 5 \cdot 10^{-10}\ \mathrm{N}^{-1}\mathrm{m}^2$. For an ideal gas, one finds from the gas law $\kappa_T = T^{-1}$ and $\kappa_p = p^{-1}$.

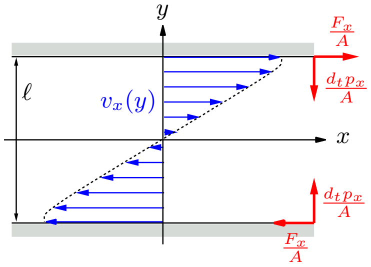


Figure 2.3.
Classical Gedankenexperiment on viscosity. Momentum is input through the movement of the plates, transported into the fluid through molecular interactions, and dissipated into heat eventually [*Landau and Lifschitz* 1981, §16]. Near the plates, the velocity field deviates from the ideal shape due to the formation of a boundary layer. Its thickness is greatly exaggerated, however.

2.2.2
Viscosity

In the classical Gedanken experiment, two large parallel plates with a fluid between them are envisaged to move past each other. Let A be the area of the plates, ℓ the distance between them, and F_x the force that acts on each of them, in opposite direction (Figure 2.3). Each of the plates inputs the momentum F_x/A per unit area and unit time into the fluid. Through molecular interactions, this momentum is transported into the adjacent fluid and on to deeper layers, setting up a shearing motion. On its way through the fluid, momentum continuously gets dissipated into heat. For the stationary case, this leads to constant velocities v_0 and $-v_0$, respectively, for the plates. For simple fluids with laminar flow and sufficiently large ℓ, one finds $F_x/A = \mu v_0/\ell$, where μ is the (dynamic) viscosity of the fluid. Instead of looking at the entire experiment, we just focus on two parallel flow planes in the fluid. In this way, we circumvent the difficulties associated with the boundary layer that inevitably develops at material boundaries. We notice that the shear stress F_x/A equals the momentum flux $\mathrm{d}_t p_x/A$ and, for two flow planes with an infinitesimal distance, we replace v_0/ℓ by the corresponding derivative. This finally yields for the momentum flux

$$\Pi_{xy} = -\mu \frac{\mathrm{d}v_x}{\mathrm{d}y} \ , \tag{2.11}$$

where x is the direction of the relative velocity between the planes and y is perpendicular to them.

The situation becomes more complicated in the multi-dimensional case where the tensorial properties of the momentum flux $\mathsf{\Pi}$ become manifest. *Landau and Lifschitz* [1981], their §15, deduce the form of $\mathsf{\Pi}$ with components Π_{ij} from the following physical postulates: (i) Friction (momentum flux) occurs only if different parts of the fluid move with different velocities. Hence, Π_{ij} depends only on spatial derivatives of the velocity, and, for sufficiently small gradients of the velocity, this dependence will be linear. (ii) There is no friction in a fluid that is in constant rotation with angular velocity $\boldsymbol{\omega}$, hence with $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{x}$. Combinations of $\partial_{x_j} v_i$ that satisfy this postulate are

$\partial_{x_j} v_i + \partial_{x_i} v_j$. The most general form of a second order tensor satisfying both postulates is

$$\Pi_{ij} = -\mu\left[\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\delta_{ij}\underbrace{\sum_k \frac{\partial v_k}{\partial x_k}}_{\nabla\cdot\mathbf{v}}\right] + \lambda\delta_{ij}\underbrace{\sum_k \frac{\partial v_k}{\partial x_k}}_{\nabla\cdot\mathbf{v}} , \tag{2.12}$$

where δ_{ij} is Kronecker's delta and the term in brackets was chosen such that its contraction is zero. Recall that the contraction corresponds to the sum over all components $i = j$ which, for a second rank tensor, yields the trace. Hence, the expression in brackets is a zero trace tensor. The coefficients μ and λ can be shown to be positive [*Landau and Lifschitz* 1981, §16] and are called coefficients of viscosity. Less formal derivations of (2.12) are given in several textbooks, e.g., in Section 5.6 of *Tritton* [1988].

For an incompressible fluid, $\nabla \cdot \mathbf{v} = 0$, hence

$$\Pi_{ij} = -\mu\left[\frac{\partial v_j}{\partial x_i} + \frac{\partial v_i}{\partial x_j}\right] . \tag{2.13}$$

This, and more generally (2.12), is Newton's law of viscosity with the dynamic viscosity μ. A fluid for which (2.12) is valid is called *Newtonian*. All gases and non-polymeric liquids belong to this class.

Parameterization of $\mu(T,p)$ The dependence of the viscosity on temperature and pressure is characteristically different for liquids and for gases. In liquids, viscosity decreases with temperature and is practically independent of pressure. It is often parameterized by

$$\frac{\mu^{\text{liq}}(T)}{\mu^{\text{liq}}(T_0)} \approx \exp\left(\beta\left[\frac{1}{T} - \frac{1}{T_0}\right]\right) . \tag{2.14}$$

Parameters for water at $T_0 = 283$ K are $\mu^{\text{liq}}(T_0) \approx 1.307 \cdot 10^{-3}$ N s m^{-2} and $\beta \approx 2.4 \cdot 10^3$ K. For water, this description is only accurate to about 10% in the temperature range from 273 K to 300 K. However, for natural systems this accuracy suffices since dissolved substances lead to higher uncertainty. Should need arise, a parameterization with four digits accuracy is given by *Cho et al.* [1999].

The viscosity of gases increases with temperature and is again practically independent of pressure. For an ideal gas, one finds

$$\mu^{\text{gas}}(T) = \sqrt{\frac{mkT}{3\sigma^2}} = \mu_0\sqrt{\frac{T}{T_0}} , \tag{2.15}$$

where m is the mass of a molecule, k is Boltzmann's constant, and σ is the molecule's effective cross sectional area. For air at $T_0 = 273$ K, $\mu_0 \approx 1.74 \cdot 10^{-5}$ N s m^{-2}.

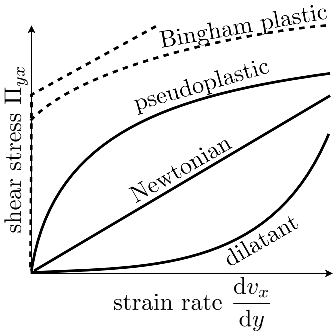


Figure 2.4.
Relation between shear stress (momentum flux) and strain rate (transverse velocity gradient) for different types of fluids. For a Newtonian fluid, the viscosity is constant whereas it increases with shear stress in a dilatant fluid and vice versa in a pseudoplastic fluid. Correspondingly, they are also referred to as shear thickening and shear thinning, respectively. The Bingham plastic may be regarded as a transition between a solid (low stress) and a fluid (high stress).

Viscous Force for Incompressible Newtonian Fluid We notice that the Π_{ij} are the components of a symmetric second order tensor $\mathsf{\Pi}$, the tensor of momentum flux. Inserting (2.13) into (2.3) and using $\nabla \cdot \mathbf{v} = 0$ leads to

$$\mathbf{F}_v = \mu \int_V [\nabla \cdot \nabla]\mathbf{v}\,\mathrm{d}V = \mu \int_V \nabla^2 \mathbf{v}\,\mathrm{d}V \ , \tag{2.16}$$

which is the force of friction for an incompressible Newtonian fluid. In Cartesian coordinates, the x-component of $\nabla^2\mathbf{v}$ is $\partial_{xx}v_x + \partial_{yy}v_x + \partial_{zz}v_x$.

Non-Newtonian Fluids Although we only consider Newtonian fluids in the following, it is worth taking a side look at other fluids. All pure fluids share the property that they cannot sustain any shear stress in static equilibrium since their constituting particles are asymptotically free. This distinguishes them from pure solids where particles are bound to their relative positions. Solids thus show elastic behavior, where an applied shear stress leads to a dislocation of particles from their equilibrium position and to a corresponding force that drives them back. Hence, a solid deforms to the point where either the two forces balance each other and the shear stress can be maintained indefinitely or the solid adjusts irreversibly, eventually breaks.

The deformation in response to an external stress (force per unit area) is called the strain and the stress-strain relation is an important material property of any solid. The analogous relation for a fluid is that between shear stress and *rate of* strain: an imposed shear stress leads to an increase of the shear velocity, hence of the transverse velocity gradient, which in turn leads to a higher momentum dissipation rate. Hence, a fluid accelerates to the point where either input flux and dissipation rate of momentum balance each other or the fluid separates thereby interrupting the transfer of momentum.

For an incompressible Newtonian fluid, (2.13) ascertains that the rate of strain is linearly related to the shear stress with viscosity μ as constant of proportionality. For non-Newtonian fluids, μ becomes a function of the shear stress Π (Figure 2.4). Depending on the shape of $\mu(\Pi)$, two types are distinguished, *dilatants*, stress-thickening fluids, and *pseudoplastics*, stress-thinning fluids. Examples of dilatants include a wide range of emulsions like cornstarch on which a person can run spectacularly but sinks in once the rapid motion is stopped. A more practical application is liquid armor in sport's or military suits. An examples of a pseudoplastic is glacier ice on which a person can easily walk, but which also flows into a valley under its own weight. Other examples are many paints which can be applied like a thin liquid with a brush but which will not drip. Many fluids are "complex" and show thinning as well as thickening regimes and even jammed states with no more flow. The latter makes the transition between fluids and solids in that the material is solid below some stress threshold and becomes fluid above. This is called a Bingham plastic. Examples of such materials include drilling mud, ketchup, or whiped cream. Since the rich macroscopic phenomenology depends on molecular-scale details, this opens the way to engineering "smart fluids" whose behavior can be controlled through electric or magnetic fields or which react to various internal variable like temperature, surfactant concentration, and emulsion density [e.g., *Brown et al.* 2010].

2.3
Flow of Incompressible Newtonian Fluid

With the results of the preceding sections, we formulate the dynamics of a uniform and incompressible Newtonian fluid and consider some particular solutions.

2.3.1
Dynamics

We choose to work in the Eulerian framework, i.e., in a coordinate system that is fixed in space. The total rate of change following the motion of a fluid element, i.e., the total derivative d_t, then consists of the local component ∂_t and of the component $\mathbf{v}\cdot\nabla$ that results from the element's velocity. Hence, $\mathrm{d}_t = \partial_t + \mathbf{v}\cdot\nabla$.

The dynamics of the fluid is obtained by combining mass balance (2.6), linear momentum balance (2.8), incompressibility assumption $\rho = \text{const}$, and viscosity model (2.16) to

$$\begin{aligned} \nabla\cdot\mathbf{v} &= 0 \\ \rho\partial_t\mathbf{v} + \rho[\mathbf{v}\cdot\nabla]\mathbf{v} &= \rho\mathbf{g} - \nabla p + \mu\nabla^2\mathbf{v} \ . \end{aligned} \tag{2.17}$$

The second equation is *Navier-Stokes'* equation for an incompressible Newtonian fluid [*Tritton* 1988]. It is often written in kinematic form, by dividing by the constant fluid density ρ and introducing the *kinematic viscosity* $\nu = \mu/\rho$, as

$$\partial_t\mathbf{v} + [\mathbf{v}\cdot\nabla]\mathbf{v} = \mathbf{g} - \frac{1}{\rho}\nabla p + \nu\nabla^2\mathbf{v} \ . \tag{2.18}$$

Before looking into solutions of these equations, we want to understand their physical meaning. The external forces may be separated into the driving forces $\rho\mathbf{g}$ and $-\nabla p$, and into the frictional, slowing, force $\mu\nabla^2\mathbf{v}$. The former may be expressed in terms of the potential

$$\psi_f := p - \rho\mathbf{g}\cdot\mathbf{x} \ , \tag{2.19}$$

which we will call the “fluid potential”. The action of the frictional force is understood best for slow motion, where $[\mathbf{v}\cdot\nabla]\mathbf{v}$ may be neglected, and in the absence of the other forces, $\partial_t\mathbf{v} = \nu\nabla^2\mathbf{v}$. With (4.7) on page 80, we recognize this as a diffusion equation for the velocity with diffusion coefficient ν [L^2/T]. Hence, the friction term dissipates the fluctuations of $\mathbf{v}$, into heat. In analogy, we recognize μ as the coefficient of momentum dissipation. Hence, we may write the momentum equation (2.17) as

$$\underbrace{\rho\mathrm{d}_t\mathbf{v}}_{\text{rate of change}} = \underbrace{-\nabla\psi_f}_{\text{generation}} + \underbrace{\mu\nabla^2\mathbf{v}}_{\text{dissipation}} \ . \tag{2.20}$$

Solving (2.18) is in general very difficult because of the nonlinear inertia term $[\mathbf{v}\cdot\nabla]\mathbf{v} = [\nabla\times\mathbf{v}]\times\mathbf{v} - \frac{1}{2}\nabla v^2$ which relates to the physical phenomenon of turbulence [*Frisch* 1995]. Short of a full analysis, which is still impossible for the general case, we may gain more insight by either approximating (2.18) to the point where analytical solutions become feasible or by doing a similarity analysis which provides scaling relations for phenomena in geometrically similar flow domains. For both approaches, it is useful to write (2.18) in dimensionless form. To that end, we consider a specific but arbitrary flow domain which is characterized by length ℓ. We denote the characteristic time for changes of the external forcing by τ and the magnitude of a typical flow velocity by u. Introducing the dimensionless variables

$$t' = \frac{t}{\tau} \ , \quad \mathbf{x}' = \frac{\mathbf{x}}{\ell} \ , \quad \mathbf{v}' = \frac{\mathbf{v}}{u} \ , \tag{2.21}$$

hence the derivatives $\partial_{t'} = \tau\partial_t$ and $\nabla' = \ell\nabla$, transforms (2.18) into

$$\frac{u}{\tau}\partial_{t'}\mathbf{v}' + \frac{u^2}{\ell}[\mathbf{v}'\cdot\nabla']\mathbf{v}' = g\widehat{\mathbf{g}} - \frac{1}{\rho\ell}\nabla' p + \frac{\nu u}{\ell^2}\nabla'^2\mathbf{v}' \ , \tag{2.22}$$

where $\widehat{\mathbf{g}} = \mathbf{g}/g$ is the unit vector pointing in the direction of $\mathbf{g}$. Notice that p remains dimensional at this point. Dividing by u^2/ℓ isolates the inertia term and leads to the dimensionless form

$$\mathrm{St}\,\partial_{t'}\mathbf{v}' + [\mathbf{v}'\cdot\nabla']\mathbf{v}' = \frac{1}{\mathrm{Fr}}\widehat{\mathbf{g}} - \nabla' p' + \frac{1}{\mathrm{Re}}\nabla'^2\mathbf{v}' \ , \tag{2.23}$$

where now $p' = p/[u^2\rho]$ is introduced and

$$\mathrm{St} := \frac{\ell}{u\tau} \ , \quad \mathrm{Fr} := \frac{u^2}{\ell g} \ , \quad \mathrm{Re} := \frac{\rho u\ell}{\mu} = \frac{u\ell}{\nu} \tag{2.24}$$

are the *Strouhal*, *Froude*, and *Reynolds number*, respectively. These dimensionless numbers quantify the relative importance of inertia with respect to that of transient forcing (St), gravity (Fr), and viscosity (Re). As will be illuminated in the following, these numbers facilitate the classification of the wide range of flow phenomena.

We first focus on the *Reynolds number* and, as a preliminary, recall that for a system with linear dynamics, different temporal and spatial frequencies do not interact, in particular they cannot exchange energy or momentum with each other. Next, recognize that, given the molecular nature of viscosity as envisaged in Section 2.2.2, dissipation is not very effective at larger scales. This is reflected by the small values of the kinematic viscosity $\nu = \mu/\rho$, the relevant property here, which is some 10^{-6} $\mathrm{m^2s^{-1}}$ for water and 10^{-5} $\mathrm{m^2s^{-1}}$ for air. One can actually show that the amplitude of a fluctuation with wavelength λ decays exponentially with time, but that the decay time is proportional to λ^2/ν (Exercise 7.1). Still, viscosity is the only process that eventually dissipates momentum, hence kinetic energy, into heat and thereby drives a fluid towards static equilibrium.

The efficiency of dissipation greatly increases in processes with a nonlinear dynamics where different modes become coupled such that momentum and energy is handed down from larger to ever smaller scales where it is eventually dissipated efficiently. Such a coupling between modes is accomplished by the term $[\mathbf{v}\cdot\nabla]\mathbf{v}$ in the Navier-Stokes equation which leads to an instability of flow patterns and causes them to break up into ever smaller units (Exercise 2.1). For the case of fully developed turbulence, for very large values of Re, this gives rise to the so-called Richardson-cascade, famously described as "*Big whorls have little whorls that feed on their velocity, and little whorls have smaller whorls and so on to viscosity*" [*Richardson* 1922]. It is also referred to as Kolmogorov's β-cascade which the K41-theory hypothesized to be self-similar [*Kolmogorov* 1941; *Frisch* 1995]. This was subsequently confirmed experimentally. Kinetic energy thus enters the fluid at some large scale through the action of gravity or some external pressure field, cascades down the so-called inertial range without loss and is eventually dissipated in the molecular range. We notice that there exist no dissipation mechanisms within the momentum cascade, hence the self-similar scaling.

The Reynolds number is thus recognized as indicator for the efficiency of macroscopic momentum dissipation. For $\mathrm{Re} \ll \mathrm{Re_{crit}}$, the inertia term $[\mathbf{v}\cdot\nabla]\mathbf{v}$ is negligible, hence also the coupling between modes, and momentum dissipation occurs at the time scale λ^2/ν. When the inertial term is not completely negligible, turbulent flow can be expected at least in parts of the flow domain. This also means that flow is transient even for stationary external forcing. The extent of turbulent regions increases with increasing magnitude of $[\mathbf{v}\cdot\nabla]\mathbf{v}$ and with it also the character of the flow changes, from laminar through intermittent to fully turbulent for $\mathrm{Re} \gg \mathrm{Re_{crit}}$. The value of $\mathrm{Re_{crit}}$ depends on the geometry of the flow domain, sometimes on small-scale properties of the boundary. For some simple domains, such critical values are reported in the engineering literature. For instance $\mathrm{Re}_{\mathrm{crit}} \approx 2'300\ldots4'000$ for a long circular pipe with diameter ℓ and with a moderately rough wall.

The *Froude number* weighs the gravitational force against the inertial force, more precisely its vertical component. Hence, for a situation with $\mathrm{Fr} \ll \mathrm{Fr}_{\mathrm{crit}}$ the vertical component of $[\mathbf{v}\cdot\nabla]\mathbf{v}$ can be neglected while the horizontal ones may remain dominating. This is in particular the case for the regime of large Reynolds numbers which then leads to turbulence that is restricted to the horizontal. This is often an excellent approximation for large-scale environmental systems including the atmosphere and the ocean. As a consequence of this, the dominant rotation of vortices, from twisters to cyclones, is in horizontal planes. We notice in passing that, depending on the perspective, there exist alternative definitions of the Froude number, in particular the square root of the definition in (2.24).

Finally, the *Strouhal number* quantifies the inertial force relative to the external forcing. It may furthermore be considered as ratio between the internal time scale ℓ/u and the external time scale τ. It is worth mentioning that an external forcing may actually force in the interior of the flow domain. The decisive point is that it is none of the fluids forces that have been incorporated into the formulation of the Navier-Stokes equation. An example are flying and swimming animals whose action on their surrounding fluid is characterized by the Strouhal number [*Taylor et al.* 2003; *Eloy* 2011]. Like for the other dimensionless numbers there also exist alternative formulations for St, in particular $\ell\omega/u$, where ω is a characteristic frequency, and, as introduced originally by *Lord Rayleigh* [1915], the inverse of the modern definition given in (2.24).

2.3.2
Approximations of Navier-Stokes Equation

We focus on slow flow with $\mathrm{Re} \ll \mathrm{Re}_{\mathrm{crit}}$, where viscosity dominates inertia, hence $[\mathbf{v}'\cdot\nabla']\mathbf{v}'$ may be neglected with respect to other terms containing $\mathbf{v}'$. With this (2.25) may be approximated by

$$\mathrm{St}\,\partial_{t'}\mathbf{v}' = \frac{1}{\mathrm{Fr}}\widehat{\mathbf{g}} - \nabla' p' + \frac{1}{\mathrm{Re}}\nabla'^2\mathbf{v}' \ . \tag{2.25}$$

This linear differential equation describes laminar flow and is referred to as the time-dependent *Stokes equation* or as the equation of creeping flow.

For many small-scale flow phenomena, in particular for fine textured porous media, external forcing varies slowly on the time scale of the internal dynamics, i.e., $\mathrm{St} \ll 1$. Flow may then be approximated as stationary,

$$\frac{1}{\mathrm{Re}}\nabla'^2\mathbf{v}' = \nabla' p' - \frac{1}{\mathrm{Fr}}\widehat{\mathbf{g}} \tag{2.26}$$

or, in dimensional form, by

$$\mu\nabla^2\mathbf{v} = \nabla p - \rho\mathbf{g} \ . \tag{2.27}$$

This is the time-independent *Stokes equation*. Notice in passing that whatever the approximation is, we can never drop $\nabla' p'$ since p is one of the state variables that must be retained in at least one term. Indeed, the $\nabla' p'$ always adjusts such that the other terms are balanced.

Example: Water Infiltrating in Soil Consider a heavy rainfall event of 1 hour duration that has 10 mm of water infiltrate a sandy soil. Hence, a characteristic time for the external forcing is $\tau = 10^3$ s. (We do not care about the factor of 3.6 since we want to get an order of magnitude estimate only, and the result is only useful if the terms are separated by several orders of magnitude.) As a characteristic length choose the radius of a large pore, $\ell = 100$ μm. To estimate a typical velocity, assume that the infiltrating water passes through 1/3 of the volume – part of the volume is occupied by solid matrix and not all pores may be conducting water – which gives $u = 30$ mm/h or about 10^{-5} m s^{-1}. (Here, we do care about the factor of 3, but for didactical reasons only, in order to emphasize that water moves at a higher velocity in the porous medium than outside.) Finally, the kinematic viscosity of water is about 10^{-6} m^{2}s^{-1}. Inserting these values into (2.24) yields

$$\mathrm{St} = 10^{-2}\ , \qquad \mathrm{Fr} = 10^{-5}\ , \qquad \mathrm{Re} = 10^{-3}\ .$$

Recalling $\mathrm{Re}_{\mathrm{crit}} = \mathcal{O}(10^3)$ for a cylindrical tube, we thus first conclude that the inertia term can be safely neglected and the Stokes approximation is valid. Next, we deduce from St that the time scale of the internal dynamics is 2 orders of magnitudes shorter than that of the external forcing. Hence, the time-independent Stokes equation (2.27) may be used. Finally, making the bold assumption $\nabla'^2 \mathbf{v}' = \mathcal{O}(1)$, we infer that gravity (Fr) dominates friction by 2 orders of magnitude. Dropping the friction term would leave us with $\nabla p = \rho \mathbf{g}$, the equation for hydrostatic equilibrium. Its solution is $p_h(\mathbf{x}) = p_0 + \rho \mathbf{g} \cdot \mathbf{x}$, with p_0 the pressure at $\mathbf{x} = 0$.

Hydrostatic equilibrium is, to a reasonable approximation, the solution of the flow problem. (This is indeed the case for many environmental systems, from porous media to the ocean and the atmosphere.) We can do quite a bit better than hydrostatic equilibrium, however. Write p as a sum of hydrostatic pressure p_h and deviation p^*, hence $p = p_h + p^*$. Insert this into (2.27) with p_h given above to find

$$\mu \nabla^2 \mathbf{v} = \nabla p^*$$

for the pressure deviation. Solutions of this equation will be obtained in Section 2.3.5. We notice that irrespective of whether $\nabla'^2 \mathbf{v}' = \mathcal{O}(1)$, the decomposition $p = p_h + p^*$ is always possible since the Stokes equation is linear.

We notice with Figure 2.5 that in the time-independent Stokes regime, the forces acting on a fluid element cancel, which means that the velocity adapts such that the viscous force (momentum dissipation) compensates the sum $\mathbf{F}_g + \mathbf{F}_p$ of gravitational and pressure force (momentum generation). In an isotropic fluid this means that $\mathbf{v}$ is always parallel to $\mathbf{F}_g + \mathbf{F}_p$.

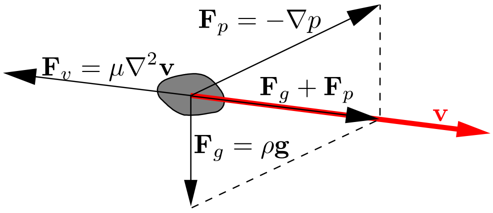


Figure 2.5. Forces on a fluid element with stationary Stokes flow. The velocity adjusts such that the resulting force vanishes. For an isotropic fluid, this leads to velocity $\mathbf{v}$ being parallel to the driving force $\mathbf{F}_g + \mathbf{F}_p$.

2.3.3 Similarity Analysis

We consider a fixed but arbitrary flow domain Ω' and for simplicity limit our interest to slow stationary flow, that is to

$$\frac{1}{\mathrm{Fr}}\widehat{\mathbf{g}} - \nabla' p' + \frac{1}{\mathrm{Re}}\nabla'^2\mathbf{v}' = 0 \;, \tag{2.28}$$

with appropriate boundary conditions. The solution of this problem may be written as

$$\mathbf{v}'(\mathbf{x}') = \mathbf{f}_v(\mathbf{x}'; \mathrm{Fr}, \mathrm{Re}) \;, \qquad p'(\mathbf{x}') = f_p(\mathbf{x}'; \mathrm{Fr}, \mathrm{Re}) \;, \tag{2.29}$$

where $\mathbf{f}_v$ and f_p are dimensionless functions, $\mathbf{f}_v$ vector-valued and f_p scalar. Depending on the geometry of the flow domain, these functions can be very complicated. Two flow problems are called *similar* if (i) their flow domains are geometrically similar, (ii) the system parameters g, ℓ, μ, and u are such that Fr and Re are the same, and (iii) the dimensionless boundary conditions are identical. Obviously, two similar flow problems lead to identical mathematical problems when expressed in dimensionless terms and their solutions may be written as

$$\frac{\mathbf{v}(\mathbf{x})}{u} = \mathbf{f}_v\big(\frac{\mathbf{x}}{\ell}; \mathrm{Fr}, \mathrm{Re}\big) \;, \qquad \frac{p(\mathbf{x})}{u^2\rho} = f_p\big(\frac{\mathbf{x}}{\ell}; \mathrm{Fr}, \mathrm{Re}\big) \;. \tag{2.30}$$

We gain further insight into the flow problem by realizing that the dimensionless solution (2.29) may be transformed with any combination of the dimensionless numbers Fr and Re. For instance, with $u = \mu\mathrm{Re}/[\rho\ell]$ from (2.24), we obtain

$$\frac{\mathbf{v}(\mathbf{x})}{\mu/[\rho\ell]} = \mathbf{f}_{v_1}\big(\frac{\mathbf{x}}{\ell}; \mathrm{Fr}, \mathrm{Re}\big) \;, \qquad \frac{p(\mathbf{x})}{\mu^2/[\rho\ell^2]} = f_{p_1}\big(\frac{\mathbf{x}}{\ell}; \mathrm{Fr}, \mathrm{Re}\big) \;, \tag{2.31}$$

with $\mathbf{f}_{v_1} = \mathrm{Re}\,\mathbf{f}_v$ and $f_{p_1} = \mathrm{Re}^2 f_p$. From this we learn that, however complicated the shape of the flow domain may be, for similar flow problems the velocity scales proportionally to μ and inversely to ℓ while pressure scales with the square of these quantities, hence proportionally to $[\mu/\ell]^2$.

In an analogous manner, we may take any quantity of interest, bring it into a convenient dimensionless form, and conclude that, besides its possible dependence on the scaled location $\mathbf{x}/\ell$, it must only depend on Fr and Re. We finally remark that this type of analysis is not limited to the Navier-Stokes equation or to stationary systems.

2.3.4
Initial and Boundary Conditions

Formulation (2.17) for the dynamics of a fluid may be interpreted as the description of the infinite ensemble of all possible evolutions for a fluid with material properties ρ and μ. A particular evolution is selected from this ensemble by specifying initial and boundary conditions. Specifying the initial condition corresponds to setting up an experiment in a chosen state or, more typically for a natural environment, to observing it. Similarly, specifying boundary conditions corresponds to forcing and experiment or at least to observe its natural forcing.

Given flow domain Ω with boundary $\partial\Omega$, a typical situation would be that $\mathbf{v}(\mathbf{x})$ and $p(\mathbf{x})$ are given throughout Ω at initial time t_0. With these functions given, the initial acceleration $\rho\partial_t\mathbf{v}$ of the fluid at each point can be calculated immediately from (2.17). Typical boundary conditions would be $\mathbf{v}(\mathbf{x})$, $p(\mathbf{x})$, or a combination of them on $\partial\Omega$. These would prescribe fluid flow across the boundary or the level of a connected reservoir.

The challenge from a practical perspective is that it may be difficult, even impossible, to setup a fluid system in a prescribed state or to observe it throughout Ω. Think of a porous medium or of the atmosphere. This means that the initial state in the physical world differs from that in the model representation. Whether this difficulty is severe depends on the flow regime. For intermediate to high Reynolds numbers, evolutions with neighboring initial states diverge rapidly such that predictions are limited to some time span. An example in case is weather prediction, the more general context is deterministic chaos [e.g., *Schuster and Just* 2005]. The situation is more benign for low Reynolds numbers where viscosity, hence dissipation, dominates. In such a system, initial states decay exponentially and evolutions from different initial states converge. Hence, a sufficient "spin-up" time suffices to bring a model that starts from an inaccurate initial state close to the real system. An example is rainfall-driven subsurface water flow.

Besides the many practical issues, boundary conditions also evoke a few fundamental problems. One of them is associated with the very definition of a boundary, another one with the representation of the relevant physical processes. To appreciate the former, consider as an example the soil-atmosphere interface while zooming down from the scale of the planetary boundary layer, a few hundred meters, all the way to the soil-water-air interface at the molecular scale of about one nanometer. To glimpse at some of the implications, we recall that understanding fluid flow in terms of the Navier-Stokes equation rests on the construct of a fluid element. Choosing its size determines which phenomena, processes, and scales are to be parameterized, i.e., represented in an averaged way as material properties, and which ones are to be described explicitly. Among others, this limits the spatial resolution of the representation.

At the large end, and for turbulent flow, choosing the size of the fluid element involves the following trade-off: The larger the fluid element, the more of the momentum cascade generated by the inertia term can be included into an effective dissipation term. As an example, within the *K41-theory*, a first-order approximation to momentum transfer in turbulent flow, the dissipation term for atmospheric flow is described by an *eddy viscosity* that is about 1 $\mathrm{m}^2\mathrm{s}^{-1}$, a factor of 10^5 larger than its molecular analog. For many flow phenomena, this may allow a non-turbulent approximation despite dominating turbulence at smaller scales. The trade-off comes from the fact that such a simplification comes at the cost of spatial resolution which then, for instance, hinders the correct estimation of the lifetime of an atmospheric low-pressure system.

At the small end, a more fundamental issue arises, fundamental in the sense of "cannot be addressed in the context of fluid dynamics". This emanates from the requirement that a fluid element is large compared to the mean distance between its constituting particles. Obviously, we cannot study small-scale processes like the movement of a contact-line where solid, water, and air meet in a triple-line, [*Thompson and Robbins* 1989] or the movement of water in clays [*Kirkpatrick et al.* 2005], since both involve interactions at distances on the order of a few particle diameters in an essential way.

This general situation is typical for environmental systems, indeed for any multi-scale system from material science through economics all the way to socio-cultural environments, and also the concepts to approach them are similar. The current credo is to get as much quantitative small-scale insight as possible and to arrive at structurally correct large-scale representation that can be parameterized at the larger scale. Ideally this involves the deduction of some scaling law from small-scale insight, which allows the projection from small to large scales, and the estimation of the appropriate parameters at the large scale. Such approximate representations of small-scale features and processes are the largest source of uncertainty for otherwise well-developed model. This is what makes models models, imperfect representations of reality that must constantly be checked against experiments.

2.3.5
Particular Solutions

Disregarding the intricacies associated with boundaries discussed above, we invoke approximations like $\mathbf{v} = 0$ at an immobile solid boundary in order to

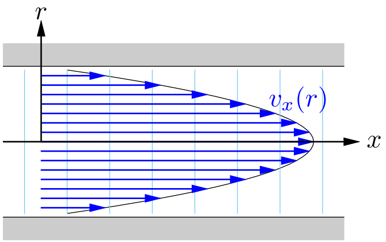


Figure 2.6. Cross-section through velocity field for *Hagen-Poiseuille* flow, i.e., slow laminar flow in long circular cylinder. The cyan lines indicate contours of the pressure field and the thin black line represents the velocity profile.

consider solutions of the Stokes equation (2.27) for some simple geometries. Focussing on the essential part, we neglect gravity, and only consider

$$\mu\nabla^2\mathbf{v} = \nabla p \; . \tag{2.32}$$

Long Cylindrical Tube Let r be the radial distance from the center and x the direction along the tube (Figure 2.6). Denote the radius of the tube by r_0. For stationary Stokes flow, the pressure gradient is constant and, due to symmetry, parallel to x. Symmetry further requires the velocity components perpendicular to x to vanish. Hence, from (2.32), $\nabla^2 v_x = \mu^{-1}\partial_x p$. Using cylinder coordinates $\{r, \varphi, x\}$ and noting that $\partial_\varphi v_x = 0 = \partial_x v_x$, we obtain $\nabla^2 v_x = r^{-1}\partial_r[r\partial_r v_x]$ and further $r^{-1}\partial_r[r\partial_r v_x] = \mu^{-1}\partial_x p$. Solving this for the boundary conditions $v_x(r_0) = 0$ and $\partial_x v_x(0) = 0$ (symmetry) yields

$$v_x(r) = -\frac{\partial_x p}{4\mu}\left[r_0^2 - r^2\right] \; , \tag{2.33}$$

the well-known parabolic velocity profile of *Hagen-Poiseuille* flow. Integrating over the cross section gives the flow

$$q = -\frac{\pi r_0^4}{8\mu}\partial_x p \tag{2.34}$$

and, dividing by the cross-sectional area πr_0^2, the flux

$$j = \frac{q}{\pi r_0^2} = -\frac{1}{8}\frac{r_0^2}{\mu}\partial_x p \tag{2.35}$$

which corresponds to the average velocity in the tube.

Bundle of Long Cylindrical Tubes Consider a densely packed bundle of cylindrical tubes, all with radius r_0, with the space between the tubes filled with some solid. With the solution for the individual tubes given in (2.33), the only notable relation is that for the average flux $\langle j\rangle$ through the “porous medium as a whole”, i.e., the flow per area of porous medium. Looking at the unit cell of the cross section, an equilateral triangle, we find $\pi/\sqrt{12}$ for the areal fraction of pores and hence for the average flux

$$\langle j\rangle = -\frac{\pi}{16\sqrt{3}}\frac{r_0^2}{\mu}\partial_x p \; . \tag{2.36}$$

Again, this may be interpreted as an average velocity, not in the tubes, however, but in the "porous medium as a whole". The average velocity in the tubes is higher by a factor of $\sqrt{12}/\pi$, i.e., by the inverse of the porosity, the volume fraction of the pores.

General Cylindrical Form For steady flow in a general cylindrical form, only the x-components of ∇p and $\mathbf{v}$ will be non-zero and $\partial_x p$ is constant. Hence, as seen above, the flow is described by $\nabla^2 v_x = \mu^{-1}\partial_x p = \text{const}$. Integration yields

$$j = -\alpha \frac{\ell^2}{\mu} \partial_x p \ , \tag{2.37}$$

where ℓ is some characteristic diameter and α is a constant that depends on the form of the cross-section. As before, we can also consider an ensemble of parallel cylindrical forms, each of them possibly with a different cross-section, and find for the average flux through such a "porous medium"

$$\langle j \rangle = -\alpha\phi \frac{\ell^2}{\mu} \partial_x p \ , \tag{2.38}$$

where ϕ is again the porosity. Choosing circular cylinders, possibly all with different radii, we arrive at the most simple, albeit not very realistic, model for a porous medium, the *parallel bundle of capillaries*.

Exercises

2.1 Momentum Cascade in Navier-Stokes Equation Study the momentum cascade introduced into the Navier-Stokes equation through the inertia term $[\mathbf{v} \cdot \nabla]\mathbf{v}$ in the simplified model $\partial_t u + \frac{1}{2}\partial_x u^2 = 0$ in a spatially infinite domain and initial state $u(x, 0) = \cos(x)$. Also discuss to what extent this is a useful approximation of the Navier-Stokes equation.

2.2 Flow through Slit Consider a horizontal infinite slit of width 2ℓ. For laminar fluid flow in this slit, calculate the velocity profile and the flow per unit length. What are the conditions for laminar flow?

2.3 Film Flow Consider a vertical, perfectly water-wet glass plate with a uniform, 20 μm thick film of water flowing down its face. Calculate the velocity profile and the flow per unit length for $T = 20°\text{C}$ and for $0°\text{C}$ (assume water to be liquid).

2.4 Compressibility of Water Consider an isothermal, vertical column of water. Let p_0 be the pressure at the surface. What is the pressure at 10 m depth?

2.5† Sap Flow in Trees The giant California redwood tree (*sequoia semperviren*) grows to heights in excess of 100 m. Discuss the possibility to transport the total transpiration flux of some 2 $\text{mm}\,\text{d}^{-1}$ through capillaries. Make reasonable assumptions whenever required.

3

Fluids in Porous Media

The dynamics of large fluid compartments – the ocean or the atmosphere, even rivers and lakes – is often dominated by the nonlinear inertia term $[\mathbf{v}\cdot\nabla]\mathbf{v}$ which causes turbulence. Boundary conditions for these compartments are simple because the direct influence of the boundaries extends only through a small fraction of the flow domain. In contrast, the dynamics of fluids in porous media is generally dominated by the porous matrix which pervades the entire volume and leads to an efficient dissipation of the fluid's kinetic energy. In return, the internal dynamics becomes rather simple, essentially described by Stokes' law. The situation becomes complicated again when multiple fluids are to be considered, as is typically the case in soils, where the air- and the water-phase coexist, or in oil reservoirs with gas-, water- and oil-phases. In the following, we first take separate looks at the architecture of porous media and at multiphase fluid systems. In the next step, we will integrate the two and move to a larger scale by averaging over the details of the pore-space.

3.1

Architecture of Porous Media

A porous medium is characterized by a partitioning of the total volume into solid matrix and pore space, with the latter being filled by one or more fluids. This partitioning indeed applies to most volume elements all the way down to some characteristic size ℓ. Since we focus on fluid movement, we are mostly interested in porous media with a connected pore space. Foams and similar porous structures will thus not be touched.

Closer scrutiny of the above characterization reveals that a precise and general definition of a porous medium is cumbersome. What is the size of ℓ, for instance? Choosing it too small would exclude stony soils while a value that is too large would accommodate strange media like karstic caverns. Whatever the eventually chosen size, ℓ has to be very much smaller than the extent of the total volume of interest. After all, there is no point in considering three grains of sand and the space between them as a porous

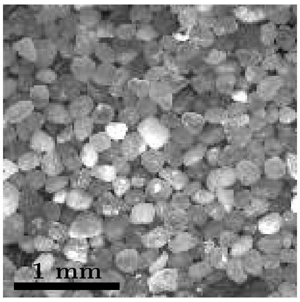


Figure 3.1.
Artificial porous medium created from sieved sand poured between two parallel glass plates, a so-called Hele-Shaw cell. This may serve as a simple two-dimensional model for a porous medium.

medium. For such objects, the shape of the flow domain must be described explicitly. An analogous difficulty arises with the connectivity of the pore space. Do, for instance, the boundaries between mineral grains in a granite constitute a connection? The answer certainly depends on the fluxes and time scale of interest. If the leakage of a mountain lake is of interest, such exceedingly thin conduits may be neglected. On the other hand, if the concern is with leakage from a nuclear depository, diffusion into grain interstices may be important.

3.1.1
Simple Porous Media

Bundles of capillaries, introduced in Section 2.3.5, would qualify as porous media and they are indeed often used to illustrate basic concepts. However, the single flow direction they permit and the constant radius of the individual pores makes them less than ideal models for soils, sediments, and rocks. A more realistic model is a heap of sand (Figure 3.1).

A simple unconsolidated medium may be described in many different ways, for instance by its grain size distribution. A significant effort has been spent to deduce the pore size distribution and hydraulic material properties from such measurements [e.g., *Arya and Paris* 1981]. The approach consists of making reasonable assumptions on the shape and packing of the grains and then, using basic geometry, to calculate the quantities of interest. For instance, assuming spherical grains with equal radii in the densest packing leads to a porosity, the volume fraction of the pore space, of $\phi = 1-\pi/[3\sqrt{2}] \approx 0.26$ (Kepler's conjecture). With this, one may calculate a characteristic radius r_c of the pore space as the ratio between its volume ϕV and its surface area A. For the densest packing with spheres of radius r, one finds $r_c = [\phi/3]r \approx 0.087r$. We notice, however, that this radius is only a rough measure: the pore size distribution is very wide, even for spherical particles all with the same radius and in a regular close-packed arrangement.

Making the grain size distribution wider reduces the porosity, since smaller spheres fill the space between larger ones. Thus also the characteristic radius r_c decreases. Natural porous media typically consist of rough and non-spherical particles. As a consequence of the friction between them, there exist

Figure 3.2.
Braided river, here beneath Haute Glacier d'Arolla, Switzerland. Many reservoirs originated from such glaciofluvial deposits. However, scales were much larger than here, typically hundreds of kilometers during interglacial periods and at the end of the last ice age.

many metastable configurations with a less dense packing. The phenomenon of metastable states becomes more pronounced for smaller particles, where surface forces are more dominant, and for oblique or platy particles. Clay soils, for instance, have a porosity that easily exceeds 0.5.

The simple porous media considered so far are dominated by a single length-scale, the characteristic radius r of the grains, or, alternatively, the corresponding radius r_c of the pores. Nature shows much more involved architectures, however, which result from a variety of generating processes.

3.1.2
Sediments

Many geomorphologic processes lead to the deposition of porous formations. Initially, and often for millions of years, these formations are unconsolidated in the sense that their matrix can only support moderate shear forces. Since the compressibility of liquids like water and oil is comparable to that of sandstone, unconsolidated formations are stable as long as the pore space is filled with some liquid. Removing it leads to compression by the weight of the overlaying material, to a reduction of the pore space, to land subsidence, and eventually even to the destruction of the reservoir.

Unconsolidated Sediments Obviously, coarse-grained deposits are of particular interest since they lead to high conductivities for fluids. Such deposits typically evolve in the headwaters of rivers, often in conjunction with nearby glaciers, where the kinetic energy of the water is still high enough to transport gravel and coarse sand. An example of such an environment is shown in Figure 3.2. The highly variable flow and deposition history of such rivers leads to a complicated architecture of the resulting reservoir. As is already apparent from the photograph, lenses of fine sand will be embedded in layers of coarse sand and gravel with occasional pockets of rubble. These may be expected to be elongated along the main direction of the depositing flow.

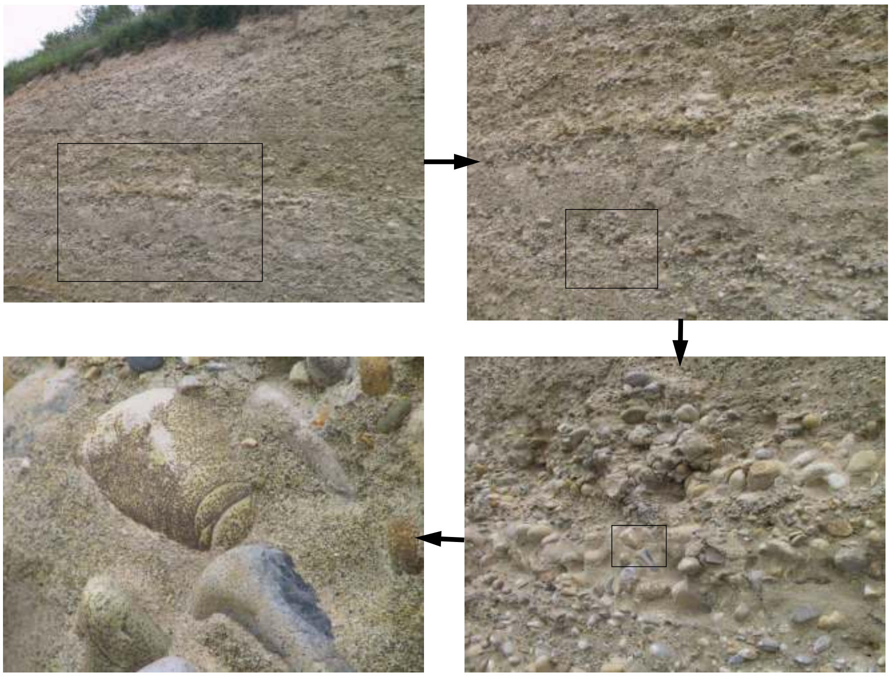


Figure 3.3. Multi-scale architecture of a typical sedimentary structure, illustrated by an outcrop at Rue de Fabron near Nice, France. Arrows indicate the path of increasing magnification and the rectangles mark the clipping for the next higher step.

Their vertical extent is rather small, however, since the flow will tend to form flat sheets due to the overwhelming impact of gravity. We thus expect exquisitely complicated, multi-scale structures in such deposits which indeed may be observed at any sand or gravel quarry or at some other outcrop (Figure 3.3). This structure will be reflected in the formations' conductivity for fluids.

Consolidated Sediments In their evolution, sediments are often consolidated and transformed through diagenetic processes. Still, they may retain a seizable fraction of their original porosity. This is for instance the case with sandstones, which are near-ideal examples of rigid porous media. Also, most igneous rocks are porous and permeable to liquids even though porosity and permeability varies greatly. Crystalline and metamorphic rocks like granite and gneiss have a very low porosity, typically less than 1%, and an exceedingly low permeability. These values may increase drastically if the rock is fractured by tectonic motion or by unloading when the overlaying material is eroded (Figure 3.4). Minerals along such fractures will slowly get dissolved by percolating water, eventually leading to a rather high permeability. This effect is even stronger with sedimentary rocks like limestone which dissolves rather

Figure 3.4. Exposed face of fractured sedimentary rock from the Mauer quarry near Heidelberg. Notice that fractures form a roughly orthogonal network which results in a strong anisotropy for the hydraulic conductivity. Height of this image is some 2 m.

quickly. Large conduits all the way to huge caverns and even subsurface streams may develop leading to so-called karstic environments. This then obviously transcends the notion of a porous medium.

3.1.3
Soils

Soils are the Earth's skin. They form at the interface between atmosphere and lithosphere from rocks and organic debris through a multitude of complicated and only partly understood processes. A still most readable and comprehensive description of soil formation is *Jenny* [1941, reprinted 1994].

Soil formation involves the physical and chemical weathering of the original rocks, i.e., the disintegration of the mineral conglomerate. Thereby some minerals are dissolved while the more stable ones are released, eventually as individual grains. Some rocks, for instance limestone, contain practically no stable minerals and over time they are dissolved almost completely. The dissolved components yield nutrients for bacteria, fungi, and eventually for plants. These in turn contribute to the soil's evolution on all scales, from the formation of organo-mineral complexes to root channels which provide a direct link between the atmosphere and deeper soil layers. Eventually, the lower life forms are followed by animals, from the miniscule enchytrae and nematodes through earthworms, typically by far the largest mass fraction of soil biota, all the way to moles and larger burrowing mammals. By digging their living space into the ground, they mix the soil and create structures with a wide range of scales. Physical processes of course continue to contribute to soil formation also after the initial weathering. These are primarily freeze-thaw cycles in cold regions, shrinking-swelling cycles in moderately humid regions, and erosion-sedimentation processes almost everywhere. The important aspect here is that the soil environment is dynamic, not static, and that

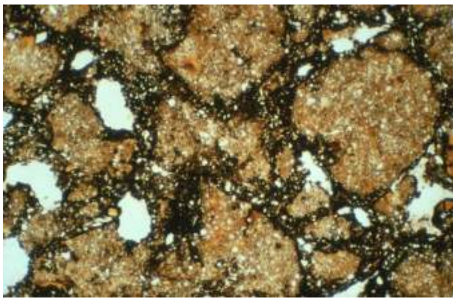


Figure 3.5.
Thin section of about 1 mm length from a loamy-clay soil. Clearly distinguishable are the system of macropores with diameters of some 0.1 mm, small soil aggregates (lighter shades of brown) with sizes of about 0.3 mm, and the system of meso- and micropores. (Image courtesy of H.-J. Vogel)

structures may be expected over a wide range of scales. While this might be reminiscent of turbulence, although at time scales of years to centuries for a typical soil layer, we notice a qualitative difference: Turbulence originates from a single "generator", the inertia term $[\mathbf{v} \cdot \nabla]\mathbf{v}$. This leads to the same structures at all scales and allows for comparatively simple scaling laws [*Frisch* 1995]. In contrast, soil structures result from many different generators and have vastly different forms. We thus cannot expect simple transitions between scales.

The multiscale nature of the pore space, which is typical for most soils, is easily appreciated by looking at even small samples as illustrated in the following examples.

Example: Meso- and Micropore System of a Loamy-Clay Soil At the smallest scale of interest here, we consider a thin section of about 1 mm length from a loamy-clay soil (Figure 3.5). The largest pores visible here have a diameter of about 0.1 mm, corresponding to a capillary height of about 0.3 m as may be calculated from (3.2). With a static water phase, these pores are thus air-filled whenever the water table is more than 0.3 m below them. In addition to the macropores, there appears a continuum of smaller pores all the way down to the resolution limit of a few micrometers. This network of so called meso- and micropores penetrates the aggregates as well as the darker material between them. Their wide range of sizes assures that some water is retained in the soil even at large distances from a water table and under rather dry conditions. This hydraulic moderation makes soils the perfect environment for many lifeforms.

Example: Macropore System of a Loamy-Clay Soil Looking at a soil sample on a larger scale, only the macropores are perceptible. Figure 3.6 shows three cross-sections, out of a stack of 120, that were obtained from a larger sample extracted from the so called B-horizon – the one below the densely rooted top-most layer with a high organic content – of a loamy-clay soil. After excavation, the sample was impregnated for stabilization and was grinded down in steps of 0.1 mm. The faces were photographed and digitized with a resolution of 0.12 mm [*Cousin et al.* 1996]. Visible structures range from the prominent cavity in the leftmost image – it is some 30 mm long, 5 mm wide, and, as inspection of images not shown here reveals, at least 2 mm deep – through almost circular voids from vertical worm channels, all the way to very small voids that originate from thin

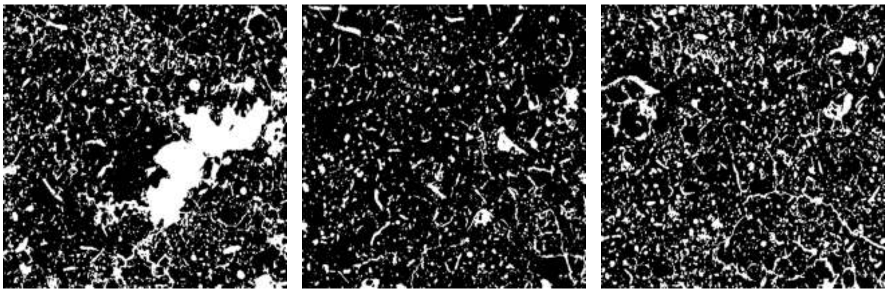


Figure 3.6. Horizontal cross-sections through a sample taken at 0.4 m depth from a loamy-clay soil near Beauce, France [*Cousin et al.* 1996]. The side length of the square sections is 48 mm with a resolution of 0.12 mm. The smallest visible pores thus are comparable to the largest pores in Figure 3.5. The vertical distance between the sections shown here is 6 mm. White represents the pore space, black the soil matrix which itself is again porous at a smaller scale. (Data courtesy of I. Cousin)

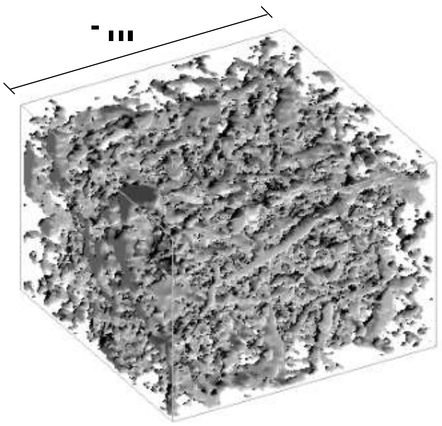


Figure 3.7. Three-dimensional reconstruction of the macropore system for a selection from the dataset shown in Figure 3.6. Resolution is 0.12 mm horizontally and 0.10 mm vertically. (Image courtesy of H.-J. Vogel)

> plant roots and enchytrae. As illustrated in Figure 3.5, much smaller pores exist in the soil matrix but they are not resolved here.
>
> A three-dimensional reconstruction of the macropore system of a sub-volume illustrates that this network of large pores is connected rather well (Figure 3.7). In a well-drained soil, macropores are only filled with water during short times after heavy rainfall events. During such events, infiltrating water bypasses the finer textured soil matrix. It is thus lost for plants and recharges groundwater directly. Thereby, it may carry solutes past the reactive surface layer and lead to a contamination of deeper soil layers or of the groundwater. Most of the time, however, the macropore system acts as a fast conduit for the exchange of gases between soil and atmosphere.

We notice in passing that the soil environment also provides a wide spectrum of physicochemical micro-sites as is apparent from the various shades of red and brown visible in Figure 3.5. They indicate the presence of organic material as well as of iron- and manganese-oxides. Obviously, this adds significantly to the complexity of the soil environment. Representing a typical soil in its full complicated glory is quite impossible and we will have to restrict ourselves to highly simplified models. To this end, we will in the following envisage a rigid porous medium with a completely water-wet matrix surface. The latter leads to a thin film of adsorbed water that covers the entire soil matrix at all times, as will be shown in the next section.

3.2
Multiple Phases

We focus on small multiphase systems with relevant scales of about a millimeter. We first realize that, despite more than two centuries of active research starting with *Young* [1805] and periodical summaries in comprehensive reviews [*Dussan* 1979; *de Gennes* 1985; *Good* 1992; *Bonn et al.* 2009], and despite the huge importance of this topic in technologies like coating, printing, and microfluidics with engineered interactions between various fluids, our quantitative understanding is still limited. Nevertheless, many aspects are understood rather well at the microscopic scale due to advances in numerical simulations of molecular dynamics and due to novel experimental techniques like isotope dilution spectroscopy [*Stiopkin et al.* 2011].

Before considering flow domains that contain multiple fluids, we shortly recall the forces that act between molecules. They all originate from the electrical fields of the molecules and may be repulsive, when molecules come so close that their electron orbitals begin to overlap, or attractive due to the interaction of permanent or induced dipoles. The forces that stem from interacting orbitals are short-ranged and decay proportional to $\exp(-r/r_0)$, where r_0 is of the order of the molecule's radius. Dipole interactions on the other hand are long-ranged, reaching across several molecules, and decay only algebraically proportional to $[r/r_0]^{-7}$. They are summarized into the *van der Waals* force which results from the *Keesom energy* for two permanent interacting dipoles, the *Debye energy* for a permanent dipole interacting with a non-polar molecule, and the *London dispersion energy* for two non-polar molecules interacting through quantum fluctuations. All these energies have an r^{-6}-dependence on distance r, hence their collection into the r^{-7} van der Waals force. Some molecules, most prominently water, form H-bonds which result from a proton (hydrogen) that is shared between two electronegative atoms (oxygen). Such bonds are considerably weaker than the corresponding covalent bonds. In water, for instance, where the molecule has a diameter of about 0.275 nm, the covalent O–H bond has the atoms at a mean distance of 0.099 nm whereas for the H-bond this is 0.117 nm. However, with the interaction potential being proportional to r^{-2}, H-bonds are much

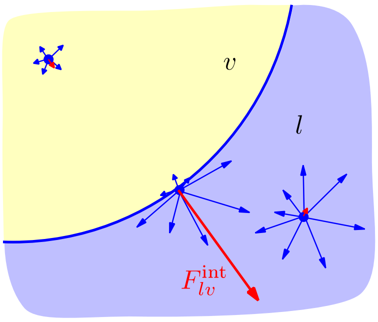


Figure 3.8.
Schematic forces between a molecule and its neighbors (blue arrows) in liquid l, vapor v, and at the interface between them. The red arrow indicates the net force on a molecule from the attraction by its neighbors. Obviously, in equilibrium it is counterbalanced by a corresponding repulsive force.
Notice that the interfacial energy depends on both of the fluids it separates.

stronger than van der Waals forces. For water, the interaction energy at room temperature is on the order of the thermal energy. As a consequence, the various bonds are not very stable. They break and form again on a time scale of a few 10^{-10} s [*Skinner* 2010].

Next, we consider the differentiation of fluids into gases and liquids. The density of gases is so low that intermolecular forces are only relevant for a small fraction of the time, during the collisions. Apart from these short intervals, molecules move freely. As a consequence, a gas fills all of the available space and all gases mix. In contrast, the density of liquids is much higher. For instance, water at room temperature is by a factor of 10^3 denser than air. Since the mass of the molecules is the same to within a factor of two – H_2O vs N_2 and O_2 – the mean distance between molecules is smaller in water by a factor of about 10. The arrangement of molecules in a liquid is determined by the equilibrium between attractive and repulsive forces under the disturbance by the thermal energy. Consequently, a liquid will maintain its density which is determined by the thermodynamic variables, and will in general not fill the entire available volume. Finally, liquids need not be miscible.

3.2.1
Interfacial Energy and Tension

Since gases mix completely, the existence of an interface demands at least one liquid. To be specific, we imagine in the following the interface between liquid l and vapor v. Nothing fundamental would change, however, if we considered two immiscible liquids like water and oil.

Forces on a molecule in the interior of a fluid average out almost perfectly over time. At an interface, this is no longer the case since some of the bonds dangle into the respective other fluid. This leads to a net normal force and to a corresponding interfacial energy (Figure 3.8). Indeed, moving a molecule from the interface to the interior releases the interfacial energy $E_{lv}^{\text{int}} = \int_0^\infty \mathbf{F}_{lv}^{\text{int}} \cdot \mathrm{d}\mathbf{x}$, where the subscripts indicate the liquid and its vapor. Integration is along the path of the molecule with the upper bound being

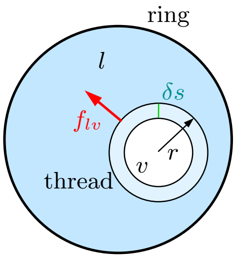


Figure 3.9.
A liquid film is suspended between an outer planar ring and a thin inner thread. The interfacial tension f_{lv} forces the thread into a circle with radius r. Reducing r by δs requires the work $2\pi r f_{lv}\delta s$. This increases the area of the liquid by $2\pi r\delta s$, hence its interfacial energy by $2\pi r\sigma_{lv}\delta s$. Since δs is arbitrary, this leads to $\sigma_{lv} = f_{lv}$. We notice in passing that in reality a film consists of two interfaces and that all the above quantities would have to be multiplied by 2. This does not affect the final relation, however.

immaterial because of the short range of the forces. Such a translocation reduces the interfacial area by A, the area taken up by the molecule. Hence we assign the areal energy density $\sigma_{lv} := E_{lv}^{\text{int}}/A$ to the interface. For the water-air interface, a thickness of some 0.3 nm has been reported [*Stiopkin et al.* 2011], about one molecular layer, and the interfacial energy density at 20°C is $\sigma_{wa} = 0.0725\ \mathrm{N\,m^{-1}}$.

In equilibrium, and with other forces negligible, the energy density associated with the interface leads to a minimal interface area for the given volumes of fluids. For two immiscible fluids, e.g., water and air, this means that one will form perfect spheres in the other. Examples are air bubbles in sirup or honey and, vice versa, free-floating water blobs sometimes demonstrated in space crafts. When other forces become important, with gravity being the most common example, the interfaces will be deformed such that the total energy of the system is minimal. An instance of this is the convenient separation of water and air in a cup of tea.

The tendency of a liquid interface to minimize its area results in tangential forces. They are for instance the culprit when a beautiful soap bubble, once punctured, collapses into an ugly drop of dirty water. Figure 3.9 illustrates that the interfacial tension, the tangential force per unit length of interface, equals the interfacial energy density,

$$\sigma_{lv} = f_{lv} \ . \tag{3.1}$$

Notice that the interfacial tension σ_{12} between fluids 1 and 2 becomes negative when the molecular attraction 1–1 is weaker than 1–2. For two liquids, this leads to the complete dissolution, e.g., of alcohol and water.

3.2.2
Discontinuity of Pressure

Tensions in a curved interface add up to a normal force which is balanced by a corresponding pressure within the fluid. Thus, pressure across a curved liquid interface is discontinuous. In order to calculate the pressure jump, first notice that every smooth surface is described locally by the two major radii of curvature, r_1 and r_2. We consider an infinitesimal area element dA which extends

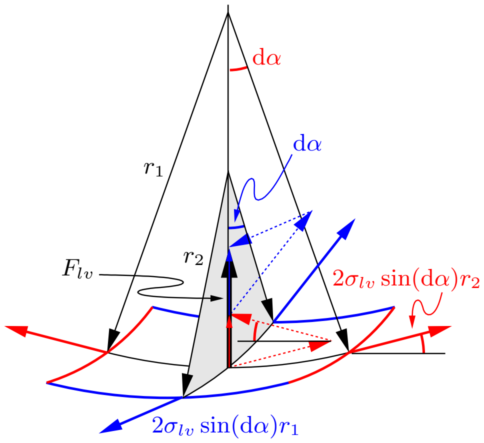


Figure 3.10. Sketch for calculating discontinuity of pressure across the interface between a liquid and its vapor. The major radii of curvature are r_1 and r_2. With half-angle $\mathrm{d}\alpha$ they span an infinitesimal area element. Tangential forces resulting from interfacial tension lead to the net normal force F_{lv} on the area element.

by the angle $\pm\mathrm{d}\alpha$ from the center of the respective radius (Figure 3.10). The net force from the interfacial tension on this element equals $F_{lv} = 4\sigma_{lv}[r_1 + r_2]\sin^2(\mathrm{d}\alpha)$. It is balanced by the pressure $F_{lv}/[4r_1r_2\sin^2(\mathrm{d}\alpha)]$. Hence

$$\Delta p_{lv} := p_l - p_v = \sigma_{lv}\left[\frac{1}{r_1} + \frac{1}{r_2}\right] = 2\sigma_{lv}H \ , \tag{3.2}$$

where

$$H = \frac{1}{2}\left[\frac{1}{r_1} + \frac{1}{r_2}\right] \tag{3.3}$$

is the *mean curvature* of the interface. Equation (3.2) is the *Young-Laplace* equation. For a spherical interface with radius r, $\Delta p_{lv} = 2\sigma_{lv}/r$.

In deducing (3.2) we used the convention that r is positive if it is located within the liquid. The two radii r_1 and r_2 need not have the same sign. This is for instance the case for so-called pendular rings that form when water is bound by capillary forces between two grains of sand (Figure 3.29). Furthermore, one or both of the radii may be infinite as is the case for a cylinder and a plane, respectively.

Further insight is gained by moving to the larger scale of the entire connected liquid. In static equilibrium and in the absence of forces other than those from the interface, the pressure at each location within the fluid is constant. If this were not the case, a flow would be induced, in contradiction to the equilibrium assumption. Hence, the pressure jump across the interface must be the same everywhere, implying that the rightmost part of (3.2) is constant. Hence, the curvature of the interface is constant. An interface between two fluids in static equilibrium is thus a minimal area surface with constant curvature.

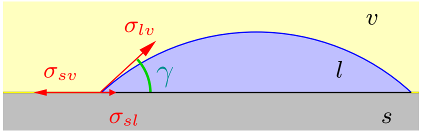


Figure 3.11.
Tensions – tangential forces per unit length – σ_{sl}, σ_{lv}, and σ_{sv} solid, a liquid, and its vapor. The contact angle γ adjust such that the horizontal components of the three forces balance each other.

3.2.3
Equilibrium Distribution of Multiple Phases

In a system with at least three phases and with all forces other than interfacial negligible, liquid interfaces form spherical calottes that are bounded by contact lines. Normal to a contact line, the interfaces extend the contact angle.

Liquid on Solid We first consider a flat solid surface in contact with a liquid and its vapor (Figure 3.11). The contact angle γ can be calculated either by minimizing the total interfacial energy, with the additional condition that the liquid volume is constant, or by balancing the horizontal components of the interfacial tensions at the contact line. The latter is typically much easier and readily leads to the *Young* equation [*Young* 1805]

$$\sigma_{lv}\cos(\gamma) = \sigma_{sv} - \sigma_{sl} \; . \tag{3.4}$$

This equation implies three regimes: (i) For $\sigma_{lv} \geq |\sigma_{sv} - \sigma_{sl}|$, the solid is partially wet, where “wet” means “covered by liquid”. An example for this is a drop of water on a car’s windshield. (ii) With increasing σ_{sl}, hence decreasing attraction of liquid molecules by the solid, the contact angle increases until, for $\sigma_{sl} > \sigma_{sv} - \sigma_{lv}$, the liquid forms a perfect sphere and the solid becomes dry. Water drops on plant leafs or on a freshly oiled wooden table are examples for this regime. (iii) With increasing σ_{sv}, the contact angle decreases until, for $\sigma_{sv} > \sigma_{sl} + \sigma_{lv}$, the liquid spreads completely and forms a films, the solid thus becoming wet as for water on a very clean glass surface.

We finally comment that all forces must balance at a stationary contact line, not just their horizontal components. Indeed, the vertical component of σ_{lv} is balanced by elastic forces in the solid, and a corresponding deformation. While the latter is minute for solids, it is observable when the solid is replaced by a more deformable substance like a gel [*Jerison et al.* 2011].

Two Liquids Imagine a pond of water with a small volume of a light oil added to its surface (Figure 3.12). Again formulate the balance of forces at the stationary contact line to obtain

$$\cos(\gamma) = \frac{\sigma_{wa}^2 - \sigma_{wo}^2 - \sigma_{oa}^2}{2\sigma_{wo}\sigma_{oa}} \; , \tag{3.5}$$

where the cosine theorem $\sigma_{wa}^2 = \sigma_{wo}^2 + \sigma_{oa}^2 + 2\sigma_{wo}\sigma_{oa}\cos(\gamma)$ was used.

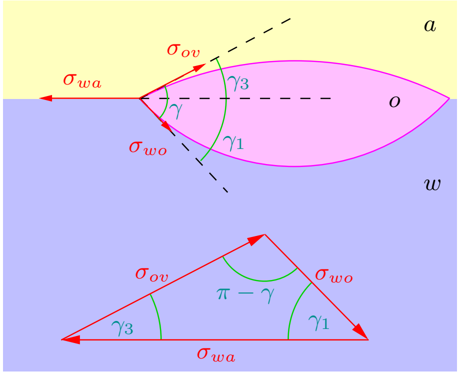


Figure 3.12.
Interfacial tensions in a water-oil-air system. The triangle shows the addition of forces. Equation (3.5) is obtained with the cosine theorem using $\cos(\pi-\gamma) = -\cos(\gamma)$.

As in the previous case, three regimes can be distinguished in a very similar way with a contact line forming if the tensions satisfy the triangle condition, i.e., any of the tensions is smaller than the sum of the other two. If this is not the case, we find either complete spreading of the oil on the water or the formation of a sphere.

Rise of Liquid in Cylindrical Capillary A liquid in a thin capillary rises to a certain height h, positive if the liquid is wetting and negative if it is non-wetting. Consider a very clean, vertical, cylindrical glass capillary with radius R whose lower end is in contact with a water table. We observe that the entire capillary is covered by a thin film of water and that a meniscus has risen to height h. This height may be calculated from the force balance at the upper end of the meniscus, where the bulk water meets the thin film extending above. Here, the force from interfacial tension, $2\pi R\sigma_{lv}$, balances the weight of the liquid column "hanging" at that line, $\pi R^2 h\rho_{lv}g$, where $\rho_{lv} = \rho_l - \rho_v$ (Figure 3.13). Neglecting the small volume of the meniscus, the capillary rise becomes

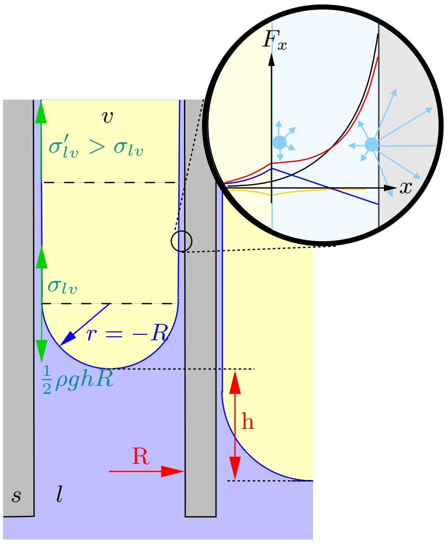
$$h = \frac{2\sigma_{lv}}{\rho_{lv}gR} = -\frac{2\sigma_{lv}}{\rho_{lv}gr} \ , \tag{3.6}$$

where R is the radius of the capillary and r the one of the interface. We notice that the water film that extends beyond the meniscus is also hold in place by interfacial forces which, apparently, are stronger than those at the meniscus. This is easily accomplished by a the liquid film becoming slightly thinner at greater height, thereby bringing the lv-interface nearer to the sl-interface such that the liquid molecules are influenced more by the stronger attraction from the solid.

For incomplete wetting, contact angle $\gamma > 0$, we notice from Figure 3.13 that only $\cos(\gamma)\sigma_{lv}$ is available for balancing the force from the weight of the liquid column. Hence, the capillary rise will be reduced to

$$h = \frac{2\cos(\gamma)\sigma_{lv}}{\rho_{lv}gR} \ . \tag{3.7}$$

Figure 3.13.
Rise of a liquid in a smooth-walled vertical capillary with radius R. Attraction of the liquid's molecules by the solid is assumed to be much larger than by the neighboring fluid, $\sigma_{12} < \sigma_{13} < 0$, hence the liquid wets the solid wall completely. Neglecting the thickness of the liquid film, the radius of the interface is $r = -R$. It is negative because it is outside of liquid 2. The dashed lines indicate imaginary cuts along which the force balance is considered. The enlargement sketches the net force (red) on a water molecule and its components from the solid (black) and from liquids 2 (blue) and 3 (yellow).



Rise of Liquid in Corrugated Capillary The equilibrium configuration of a liquid in a smooth-walled capillary with constant radius R is unique, i.e., independent of the path on which the equilibrium was established. This is not true anymore for a corrugated capillary since the condition for equilibrium is satisfied for different heights of the liquid column. For a completely wetting liquid and a capillary whose cross-section is circular at every height h equilibrium is still given by (3.6), although with a possibly varying radius $R(h)$. Such an equilibrium is stable if the gain in gravitational energy by an infinitesimal raise of the meniscus by $\mathrm{d}h$ would be larger than the corresponding reduction in interfacial energy, hence if

$$\frac{\mathrm{d}R}{\mathrm{d}h} > -\frac{\rho_{lv} g}{2\sigma_{lv}} R^2 \tag{3.8}$$

at the equilibrium height (Figure 3.14). Which one of the stable equilibria is attained eventually depends on the history.

The situation becomes even more complicated if, instead of a single capillary, we consider a three-dimensional network of corrugated pores. There we expect many possible equilibrium configurations of the liquid phase for any given position of the liquid table. As a starter, consider two corrugated, roughly parallel vertical capillaries that are connected by a horizontal capillary. Depending on details of the geometry of this structure, the fluid phases may be continuous – there is a continuous path between any two points in a phase that never leaves it – or any of them may be discontinuous with entrapped vapor or liquid. In such a medium, the actually realized configuration does not only depend on the previous configuration, but also on the speed with which the boundary conditions are changed.

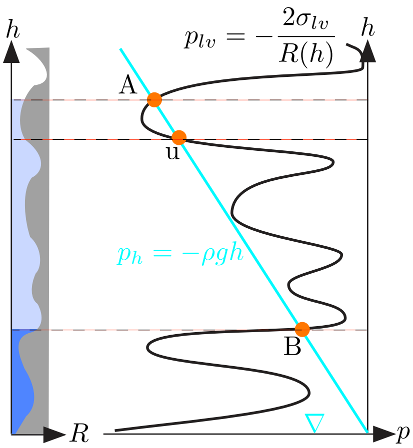


Figure 3.14.
Corrugated capillary with circular cross-section with radius $R(h)$ (left). With the liquid table at the lower end, there exist two stable configurations for the liquid – A (light blue) and B (dark blue) – with the height of the respective menisci determined by the balance between capillary pressure p_{lv} and hydrostatic pressure p_h (right). There is a further equilibrium point u which is unstable, however. A small deviation will grow until one of the stable points is reached. Depending on the geometry, there may exist more, or fewer, stable as well as unstable equilibrium points. (Adapted from *Haines* [1930].)

Capillary Length We notice that (3.7) yields a convenient measure for the length scale below which capillarity must be taken into account. If we deem capillary effects as important when the height of rise exceeds the diameter of the tube, then we find the characteristic *capillary length*

$$\ell_c = \sqrt{\frac{|\cos(\gamma)|\sigma}{\rho g}} . \tag{3.9}$$

For a water-air system in a perfectly wettable medium with $\gamma = 0$ one finds $\ell_c \approx 2.7$ mm. Replacing the air by an oil whose density is 1% less than that of water and with $\sigma_{wo} = 50 \text{ mN m}^{-1}$ leads to a capillary length that is about an order of magnitude larger, $\ell_c \approx 22$ mm.

3.3
Transition to Continuum Scale

The detailed description of the pore geometry, with relevant lengths in the micrometer range, and of pore-scale phenomena is not feasible for any domain of practical environmental interest. These have extents on the order of 1 m, often far more. Even for just 1 m^3, we would have to cope with some 10^{18} degrees of freedom! Obviously, we must come up with an averaged description along a similar line as the one that leads from gas dynamics to thermodynamics. Hence, we want to make the transition from the discrete pore space to the continuum and to macroscopic field variables which describe the observed phenomena.

3.3.1
Representative Elementary Volume

In a pore-scale representation, each point belongs to one of the phases, either to the solid matrix or to some fluid. For a formal representation, we use the *indicator function* which for phase i is defined as

$$\chi_i(\mathbf{x}) = \begin{cases} 1 & , \ \mathbf{x} \in \text{phase } i, \\ 0 & , \ \mathbf{x} \notin \text{phase } i. \end{cases} \tag{3.10}$$

Clearly, such a definition is not very useful operationally, since the decision of whether a point belongs to i or not is difficult and may even be impossible, e.g., in a fractal medium. However, this difficulty may be circumvented by thinking of a regularized representation of the pore space, i.e., one where the true pore space has been convolved with a sufficiently narrow test function. Such a regularization is done implicitly by all measuring instruments. An example of a measured indicator function of the pore space is shown in Figure 3.6, provided that white represents 1 and black 0.

A natural way for reducing the pore-scale description to a more manageable form is the averaging of the indicator functions with some convenient weight function $\kappa(\mathbf{x})$ to obtain

$$\theta_i(\mathbf{x}) := \langle\chi_i\rangle(\mathbf{x}) := \int_\Omega \chi_i(\mathbf{x}+\mathbf{x}')\kappa(\mathbf{x}')\,\mathrm{d}\mathbf{x}' \ , \tag{3.11}$$

where Ω is the support of κ and θ_i is the volume fraction of phase i, sometimes also called the macroscopic phase density. Similarly, we obtain for the average of some microscopic property α^μ of phase i the macroscopic value

$$\alpha^{\mathrm{m}}(\mathbf{x}) := \langle\alpha^\mu\chi_i\rangle(\mathbf{x}) := \int_\Omega \alpha^\mu(\mathbf{x}+\mathbf{x}')\chi_i(\mathbf{x}+\mathbf{x}')\kappa(\mathbf{x}')\,\mathrm{d}\mathbf{x}' \ . \tag{3.12}$$

For such averages to be unbiased, we require $\kappa(\mathbf{x}) \geq 0$ and $\int_\Omega \kappa(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$. Popular choices for κ are a (i) constant within a spherical support of radius σ and 0 outside or (ii) the n-dimensional Gaussian $[2\pi\sigma^2]^{n/2}\exp(-|\mathbf{x}|^2/[2\sigma^2])$ with zero mean and variance σ^2 which averages over a typical distance σ. The latter has the advantage that the averages are smooth, i.e., their spatial derivatives of arbitrary order are continuous.

We notice that averaging transforms the discrete microscopic description, where each point belongs to exactly one phase, to a continuous representation where at each point the densities of the macroscopic quantities are given (Figure 3.15). Hence the notion of a transition to the continuum scale.

In general, definition (3.11) depends on the choice of κ which makes all subsequent considerations rather cumbersome. However, if the indicator function χ_i is sufficiently uniform at some large scale, increasing the averaging volume Ω will eventually lead to a value for θ_i that depends neither on

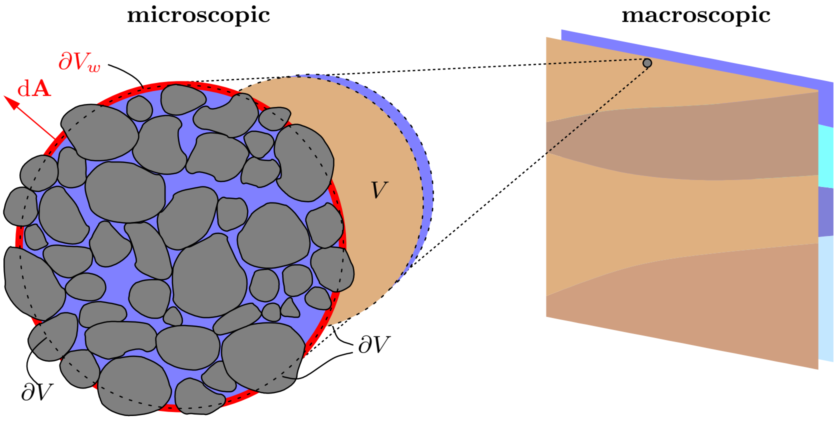


Figure 3.15. Transition from pore-scale (microscopic) to continuum (macroscopic) representation. Consider a macroscopic volume V with boundary ∂V (dotted line). Microscopically, the detailed distribution of all the phases is available, e.g., of the water phase $V_w \subset V$ with external boundary $\partial V_w \subset \partial V$ (red line). Macroscopically, the phases and possibly other quantities are replaced by the superposition of continuous fields (uniformly colored regions). These fields may vary in space, but on a much larger scale than that of the averaging volume.

the size nor on the form of Ω. Such an averaging volume is called a *representative elementary volume*. It is the means for an objective transition from a pore-scale representation to one at the continuum scale and is thus of fundamental importance. Early studies of porous media considered averaging functions that were constant within some bounded support and 0 outside, and introduced the notion of a *representative elementary volume* (REV) for a volume of linear extent ℓ [*Bear* 1972].

Example: REV for Porosity of Macropore System We consider the volume fraction of the pore space, the so called *porosity*, of the soil sample illustrated in Figure 3.6 and Figure 3.7 and use a cube of side length ℓ as averaging volume with a weight function that is constant within the cube. From a coverage of the entire sample volume with such cubes, we obtain the distribution of local porosities as a function of ℓ (Figure 3.16). For very small values of ℓ, this distribution is very broad with extremes given by cubes that are completely within the matrix or in the pore space. With increasing values of ℓ, the representation of both, matrix and pore space, become coarser and the difference between cubes decreases. Eventually, we may postulate an REV with a size of about 17 mm. Fluctuations at larger scales would then be attributed to macroscopic variability.

In the following, we will use "REV" as shorthand for the requirement that some microscopic quantity becomes approximately constant when averaged with a weight function of characteristic extent ℓ, even if this function is smooth and possibly with unbounded support.

We realize that the transition to continuous fields replaces the original pore-scale representation by a new one. Considering some process like water

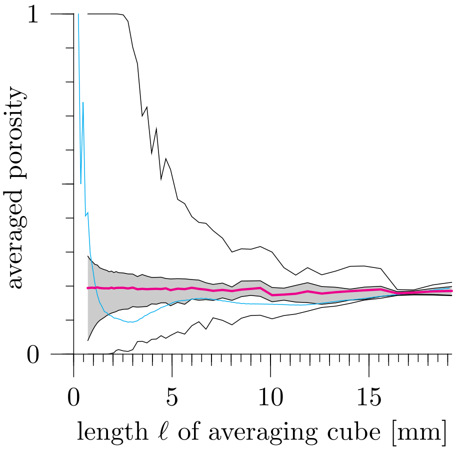


Figure 3.16.
Estimated porosity of soil sample from Figure 3.6 as a function of averaging cube's length. The cyan curve represents a particular location. The other curves represent the ensemble of all cubes: average (magenta), minimum and maximum, and the two quartiles. Half of all values are within the gray band. The linear extent of a reasonable REV would be some 17 mm.

flow or heat transport in both representations, we may hope that the respective descriptions approach each other at a sufficiently large scale, ideally that of the REV. However, large deviations between the two are to be expected at smaller scales. Hence, despite the fact that we will eventually formulate differential equations for the dynamics of transport processes in porous media, evaluating the ensuing solutions at scales smaller than the REV will generally not reflect the physical processes correctly. This becomes particularly virulent near boundaries. This issue is of course not limited to the transition from the pore-scale to the continuum, but applies to all transitions between scales that involve some kind of averaging and also to measuring processes, where the measuring volume takes the role of the REV.

3.3.2
Texture

In soil science, “texture” refers to the grain size distribution of a soil while “structure” refers to the soil's aggregation into crumbs and blocks. Here, we employ a more general semantic with *“texture”* encompassing the small-scale shapes of a porous medium for which a statistical representation suffices. *“Structure”* on the other hand refers to features that must be represented explicitly. Consider for instance a soil at a scale of 1 m. Soil layers or earthworm channels here would belong to the structure while the system of meso- and micropores would constitute the texture. This extended notion is useful at any scale and in particular for hierarchically heterogeneous media.

After the transition from the pore scale to the continuum scale, the texture of the pore space is reduced to a few statistical quantities like volume and surface densities, correlation functions and lengths, and possibly some connectivity functions. Here, we only consider the lowest order description, the volume density of the pore space which is usually referred to as the *porosity* ϕ.

Formally, $\phi = \langle \chi_\phi \rangle$ is obtained from the indicator function χ_ϕ of the pore space through (3.11).

It is often convenient to also introduce the *bulk density* ρ_b which is obtained by averaging the mass density ρ_m of the solid matrix. With (3.12) this may be written as

$$\rho_b = \big\langle \rho_m(\mathbf{x})[1 - \chi_\phi(\mathbf{x})] \big\rangle \; . \tag{3.13}$$

Assuming ρ_m to be constant, which is a good approximation for mineral porous media for which $\rho_m \approx 2.65 \cdot 10^3$ kg m^{-3}, we obtain

$$\rho_b = \rho_m [1 - \phi] \; . \tag{3.14}$$

Since both ρ_b and ρ_m can be measured with reasonable effort for many porous media, this equality is often used to calculate the porosity ϕ.

3.3.3
State Variables

To describe the macroscopic state of a fluid in a porous medium in thermodynamic equilibrium, the amount of fluid and its potential energy must be given. We will find later, in Section 3.4, that these two state variables may be linked with each other, although through a rather complicated hysteric relation.

Amount of Fluid The volume fraction θ_i of phase i – the volume of phase i per volume of porous medium – was already defined in (3.11). Sometimes, the *saturation*

$$\Theta_i := \frac{\theta_i}{\phi} \tag{3.15}$$

is used instead of the volume fraction. It most useful for dealing with incompressible media, which will be the case in the following.

In analogy to the bulk density (3.14), the macroscopic mass density of phase i could be introduced. We will not do this, however, and will always write it explicitly as $\rho_i \theta_i$, where ρ_i is the mass density of the pure fluid.

Potential Energy Potentials are more convenient than forces for describing fluid flow through porous media. We define the potential energy, more precisely its density, by the energy required to move a unit volume of fluid from some reference state to a particular state in the porous medium. In general, we define the state of an element of fluid i by its height z, its pressure p_i and temperature T_i, and by the concentrations C_{ij} of chemical species j. The energy of this state is referred to that in a reference state which we choose as $z = z_0$, $p = p_0$, $T = T_0$, and $C_{ij} = 0$.

In the following, we consider the simple case of isothermal flow of a pure fluid. The only contributions to the potential energy ψ_i per unit volume of fluid i then come from gravity and from the pressure difference. Hence,

$$\psi_i(\mathbf{x}) = \underbrace{p_i(\mathbf{x}) - p_0}_{\psi_p} \underbrace{- \int_{z_0}^{z} \rho_i(z')g \, \mathrm{d}z'}_{\psi_g} , \tag{3.16}$$

where z is the downward pointing component of $\mathbf{x}$, ψ_p is called the *pressure potential* and ψ_g the *gravity potential.* For an incompressible medium, this reduces to

$$\psi_i(\mathbf{x}) = [p_i(\mathbf{x}) - p_0] - [z - z_0]\rho_i g , \tag{3.17}$$

which is analogous to the fluid potential introduced in (2.19). The pressure term p_i may consist of several components and actually be of a rather complicated nature. To appreciate this, we start out from the most simple case and proceed to more complicated ones.

First consider the two fluids water and air in a water-wet, rigid porous medium and choose the ambient air pressure as p_0. Since the density of air is three orders of magnitudes smaller than that of water, the hydrostatic pressure in the air is negligible. For bound water, $p_w < p_0$, the only contribution to the pressure potential comes from the interfacial forces and is given by (3.2). This is the typical situation in a well-aerated soil at a sufficient height above the water table. Things get a bit more complicated when the air is under pressure $p_a > p_0$. This occurs for instance when an irrigation front in an initially unsaturated soil compresses the air ahead or in a gas reservoir. Then, $p_a - p_0$ gets added to the interfacial pressure jump. Similarly, when the density of some overlying fluid i is not negligible, the hydrostatic pressure $\rho_i gh$, where h is the vertical extent of fluid i, has to be added.

For free water in the above porous medium, $p_w \geq p_0$. Now, water is not kept in place by interfacial forces anymore and the hydrostatic pressure $p_w = \rho_w gz$ develops, where z is the depth below the corresponding free water table. This is the typical situation in groundwater. A note in passing: bound and free states are sometimes, and in a somewhat sloppy manner, associated with an unsaturated and saturated medium, respectively. The difference between the two notions becomes apparent when we consider the capillary in Figure 3.13. For a height less than h, the capillary is saturated with fluid 2. However, the pressure in this region is lower than the pressure in fluid 3, hence fluid 2 is bound. In a porous medium, this zone of full saturation, but still bound fluid, belongs to the *capillary fringe.*

The situation becomes considerable more complicated with saturated compressible porous media. Here, the pressure in the fluid phase compensates a fraction of the weight of the overlying mass. Then, the hydrostatic pressure does not only depend on ρ_i but also on the density of the matrix and on the relative compressibilities of fluid and matrix. This component is sometimes referred to as *overload* pressure or overload potential.

Finally, we consider the case of a medium that dries out ever more and again assume a water-wet rigid porous medium. As water is removed, the largest pores empty first because water is bound less tightly there as may be seen from (3.2). Assume circular and cylindrical pores for simplicity and rearrange (3.2) for the water-air system as $p_w = p_a + 2\sigma_{wa}/r$, where $r < 0$ is the radius of the interface. With $p_a = 10^5$ Pa and $\sigma_{wa} = 0.0725$ N m^{-2}, we find that p_w would become negative once the interfacial radius falls below 1.45 μm. It is then appropriate to replace the notion of a pressure by that of an energy density, actually by the chemical potential of the fluid phase. We will thus write (3.17) as

$$\psi_i(\mathbf{x}) = \psi_{i_m} - [z - z_0]\rho_i g \tag{3.18}$$

and refer to ψ_{i_m} as the *matric potential*. In our notation, this potential is negative when the fluid is bound and positive when it is free. For the case of water – $i \to w$ –, ψ_w is usually called the *soil water potential* and the subscript w is dropped in the matric potential

$$\psi_m = p_w - p_a \ . \tag{3.19}$$

A thermodynamic definition of ψ_m is postponed to Section 8.1.2.

Head It is sometimes convenient to express the potential energy per unit weight $\rho_i g$ instead of unit volume. Then, (3.18) becomes

$$h_i(\mathbf{x}) = h_{i_m} - [z - z_0] \ , \tag{3.20}$$

where $h_i = \psi_i/[\rho_i g]$ is the *hydraulic head* and h_{i_m} the *matric head*. The interpretation of these quantities is that the energy density is expressed in terms of the height of an equivalent fluid column. Notice that a negative value of the matric head indicates bound water with the meaning that the corresponding potential could hold a hanging fluid column of the given height.

3.3.4
Mass Balance

We first envisage a macroscopically uniform and rigid porous medium whose pore space is completely filled with a single fluid, say water of mass density ρ_w. We consider a macroscopic volume element V with boundary ∂V. This element is chosen to be an REV and its shape shall be arbitrary but invariable (Eulerian perspective). Let $V_w \subset V$ be the water-filled part and let $\partial V_w \subset \partial V$ be its external boundary (Figure 3.15). Conservation of the mass of water means that the rate of change of the mass within a volume equals the mass flow across the volume's boundary. Microscopically, the volume to be considered is V_w with its boundary ∂V_w. From the macroscopic perspective, however, there exists no explicit representation of water phase and solid matrix anymore and we need to operate with V and ∂V, respectively. Formulating the microscopic conservation of the mass of water yields

$$\partial_t \int_{V_w} \rho_w^\mu \,\mathrm{d}V = -\int_{\partial V_w} \rho_w^\mu \mathbf{v}^\mu \cdot \mathrm{d}\mathbf{A} = -\int_{V_w} \nabla \cdot [\rho_w^\mu \mathbf{v}^\mu] \,\mathrm{d}V \ , \tag{3.21}$$

where the superscript μ indicates microscale quantities, the minus sign comes from the area element $\mathrm{d}\mathbf{A}$ pointing outwards, and Gauss' theorem has been applied for the second equality. Notice that in writing (3.21), we have implicitly used $\mathbf{v} = 0|_{\partial V \setminus \partial V_w}$ and $\rho_w = 0|_{V \setminus V_w}$ in order to apply Gauss' theorem. In the next step, we use that divergence and integral may be interchanged in the last term (Exercise 3.11) and obtain

$$\partial_t \underbrace{\int_{V_w} \rho_w^\mu \,\mathrm{d}V}_{\|V_w\| \langle \rho_w^\mu \rangle_w} + \nabla \cdot \underbrace{\int_{V_w} \rho_w^\mu \mathbf{v}^\mu \,\mathrm{d}V}_{\|V_w\| \langle \rho_w^\mu \mathbf{v}^\mu \rangle_w} = 0 \ , \tag{3.22}$$

where $\|V_w\|$ is the volume of V_w and $\langle \dots \rangle_w$ indicates averaging over V_w. Dividing this by $\|V\|$, and noting that $\|V_w\|/\|V\|$ equals the volume fraction θ_w of the water phase, yields

$$\partial_t [\theta_w \langle \rho_w^\mu \rangle_w] + \nabla \cdot [\theta_w \langle \rho_w^\mu \mathbf{v}^\mu \rangle_w] = 0 \ . \tag{3.23}$$

This is the conservation of mass formulated at the macroscopic scale. We may identify the terms in brackets as the macroscopic mass density and mass flux, respectively. The second term is difficult since it depends on the microscopic correlation between density and velocity. This is not severe for the case of weakly compressible fluid, however. For this case, the correlation is small because ρ_w^μ depends only weakly on *pressure* whereas $\mathbf{v}^\mu$ depends on its *gradient*. Hence, $\langle \rho_w^\mu \mathbf{v}^\mu \rangle_w = \langle \rho_w^\mu \rangle_w \langle \mathbf{v}^\mu \rangle_w$ is a good approximation.

We choose to finally formulate the mass balance in terms of the real density of water and of the macroscopic volume flux

$$\rho_w := \langle \rho_w^\mu \rangle_w \qquad \text{and} \qquad \mathbf{j}_w := \theta_w \langle \mathbf{v}^\mu \rangle_w \ , \tag{3.24}$$

respectively, where superscripts are dropped in favor of a more compact notation. We thus obtain

$$\partial_t [\theta_w \rho_w] + \nabla \cdot [\rho_w \mathbf{j}_w] = 0 \ . \tag{3.25}$$

Up to now, we only considered exchange of water across some boundary ∂V. We will in the following also look at situations, where water is extracted directly out of V, for instance by pumping wells or plant roots. Let γ_w be the volume extraction rate, that is the volume of water that is extracted per unit time and unit volume. Since this formulation is already at the macroscopic scale, it is easily accommodated by modifying (3.25) into

$$\partial_t [\theta_w \rho_w] + \nabla \cdot [\rho_w \mathbf{j}_w] = -\rho_w \gamma_w \ . \tag{3.26}$$

Finally, we comment that (3.25)–(3.26) may be written down directly as conservation equations at the macroscopic scale. However, while formally correct – of course: it is just a mathematical statement – this is not so useful because the quantities ρ_w and $\mathbf{j}_w$ could then not be related to microscopic quantities and pure material properties as is the case with (3.24).

3.3.5
Empirical Flux Law

Macroscopic conservation of linear momentum could be formulated in analogy to the conservation of mass by averaging the microscopic momentum equation [*Gray and Hassanizadeh* 1998]. We will follow a different path, however, and rely on qualitative arguments – explicit calculation as was done for the simple geometries of Section 2.3.5 are no longer feasible – to arrive at a macroscopic flux law. Thereby, all the complications arising from the intricate geometry of a porous medium are hidden in appropriate material properties.

Stationary Newtonian Flow We assume that (i) external forcing changes so slowly that the flow may be considered as stationary, (ii) flow is so slow that inertia may be neglected, and (iii) the extent of the flow domain is so small that gravity can be neglected, relative to friction. These conditions can be quantified through the dimensionless numbers defined in (2.24). For instance, (iii) may be formulated as $\mathrm{Fr} \gg \mathrm{Re}$, or, using (2.24), as $\ell^2 g/[u\nu] \ll 1$. These assumptions lead to the simplified Stokes equation (2.32), $\mu\nabla^2\mathbf{v} = \nabla p$. We notice that this is a linear equation, i.e., if $\{\mathbf{v}, p\}$ is a solution then $\{\alpha\mathbf{v}, \alpha p\}$ is also one. We further recall that $\mathbf{v}$ is parallel to $-\nabla^2\mathbf{v}$ (Figure 2.5), hence $\mathbf{v} = -\kappa\nabla^2\mathbf{v}$ with $\kappa(\mathbf{x})$ some scalar function, typically highly variable in space, which represents the complicated pore geometry. Notice that $[\kappa] = \mathrm{L}^2$. With this, we obtain the generalization

$$\mathbf{v} = -\frac{\kappa(\mathbf{x})}{\mu}\nabla p \tag{3.27}$$

of Hagen-Poiseuille's law. It establishes a linear relation between $\mathbf{v}$ and $-\nabla p$ with a constant of proportionality that scales with the square of a length, a typical diameter in the pore space, and inversely with the fluid's viscosity.

After deriving the flux law at the pore scale, we deduce the corresponding relation at the continuum scale by averaging over a plane REV. Denote this plane by Γ, its orthonormal vector by $\widehat{\mathbf{n}}$, and the intersection of Γ with the pore space by Γ_w (Figure 3.17). The macroscopic volume flow through Γ_w is then given by

$$\|\Gamma\|\mathbf{j}_w \cdot \widehat{\mathbf{n}} = \int_{\Gamma_w} \mathbf{v}^\mu \cdot \widehat{\mathbf{n}}\,\mathrm{d}A = \|\Gamma_w\|\langle\mathbf{v}^\mu \cdot \widehat{\mathbf{n}}\rangle_w\ , \tag{3.28}$$

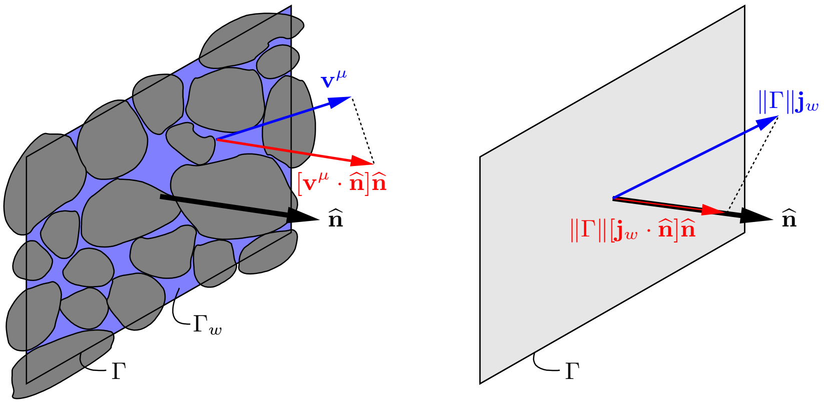


Figure 3.17. Flow through macroscopic planar REV Γ with orthonormal vector $\widehat{\mathbf{n}}$ in microscopic (**left**) and macroscopic (**right**) representation, respectively. The red arrows indicate the projection of velocity or flow vectors (blue) into the direction of $\widehat{\mathbf{n}}$. $\Gamma_w \subset \Gamma$ is the region occupied by the water (light blue).

where $\|\Gamma_w\|$ is the area of Γ_w and $\langle \dots \rangle_w$ denotes the average over this area. Next, we divide by $\|\Gamma\|$ and notice that $\|\Gamma_w\|/\|\Gamma\| = \|V_w\|/\|V\| = \theta_w$ if Γ is an REV and if the pore space is sufficiently irregular. Since $\widehat{\mathbf{n}}$ is arbitrary, this again leads to $\mathbf{j}_w = \theta_w \langle \mathbf{v}^\mu \rangle_w$. Inserting (3.27) finally produces

$$\mathbf{j}_w = -\theta_w \frac{1}{\mu} \langle \kappa \nabla p^\mu \rangle_w \ . \tag{3.29}$$

Notice that κ is not constant but varies in space at least as rapidly as ∇p^μ. Hence, the average depends strongly on correlations between κ and ∇p^μ, i.e., on details of the pore space. We resolve this difficulty by (i) recalling that $\kappa(\mathbf{x})$ represents the geometry of the pore space and neither depends on $\mathbf{v}$ nor on ∇p and by (ii) invoking the linearity of Stokes law. Hence, the magnitude of the microscopic pressure gradient is proportional to that of the macroscopic one, $\|\nabla p^\mu\| \propto \|\nabla p\|$, even though they are almost never parallel due to the geometry of the pore space. With this, $\nabla p^\mu = \mathsf{a}(\mathbf{x})\nabla p$ and

$$\mathbf{j}_w = -\frac{1}{\mu} \underbrace{\theta_w \langle \kappa \mathsf{a} \rangle_w}_{=:\mathsf{k}} \nabla p = -\frac{\mathsf{k}}{\mu} \nabla p \ , \tag{3.30}$$

where k, the *permeability*, is a second rank, symmetric tensor. This law is named after *Darcy* who, while studying water flow through sand filters in order to optimize the water supplies of the city of Dijon, suggested a linear relation between the water flow and the height difference between the water levels at the inlet and outlet ends of the columns [*Darcy* 1856].

The permeability k is a material property of the pore space. For microscopically anisotropic media, k is also anisotropic and causes the flux direction to deviate from that of its driving force $-\nabla p$ (Figure 3.18). For isotropic

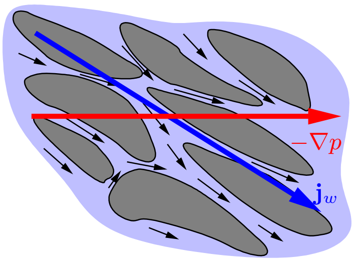


Figure 3.18.
The macroscopic flux $\mathbf{j}_w$ and its driving force $-\nabla p$ (large arrows) in anisotropic media need not be parallel, even though their microscopic counterparts $\mathbf{v}^\mu$ and $-\nabla p^\mu$ are (small arrows).

media, k may be replaced by a scalar. Finally, notice that the elements of k scale like ℓ^2, where ℓ is a characteristic diameter of the pore space.

Sometimes, particularly in the engineering literature, the elements of k are given in units of *darcy* with 1 darcy $= 0.987 \cdot 10^{-12}$ m^{-2}.

Finally notice that the volume flux $\mathbf{j}_w$ may be interpreted as a velocity, it is sometimes referred to as *Darcy velocity*, and corresponds to the velocity that would be required in the absence of the porous matrix, to sustain the given volume flux. It is related to the average microscopic velocity $\langle \mathbf{v}^\mu \rangle_w$ by (3.24). This average is usually called the *pore water velocity* and, in formulations at the macroscopic scale, denoted by $\mathbf{v}$. Hence,

$$\mathbf{j}_w = \theta_w \mathbf{v} \ . \tag{3.31}$$

The two constants that occur in Darcy's law, k and μ, are often merged into one, the *hydraulic conductivity*

$$\mathsf{K} := \frac{\mathsf{k}}{\mu} \ , \tag{3.32}$$

which may be interpreted as a material property of some geologic environment, for instance of an aquifer or an oil reservoir. Obviously, such a definition is only useful when the fluid stays always the same.

Multiple Driving Forces We deduced Darcy's law under the premiss of stationary and slow flow of a Newtonian fluid that is only driven by a pressure gradient. Extension to gravity as an additional driving force is straightforward whereas releasing the other assumptions leads to a much more complicated analysis and may even invalidate the very concept of relating macroscale velocity and macroscale driving force.

When there are driving forces other then the negative pressure gradient, we can include them in our analysis by simply replacing $-\nabla p$ with the sum $\mathbf{F}$ of all driving forces per unit volume. This is correct for an isotropic fluid and as long as the flow may be assumed to be stationary (Figure 2.5). The case encountered most often is a fluid driven by gravity and pressure gradient for which Darcy's law may be written as

$$\mathbf{j}_w = -\frac{1}{\mu}\mathsf{k}[\nabla p - \rho \mathbf{g}] \ . \tag{3.33}$$

We notice in passing that static equilibrium, that is $\mathbf{j}_w = 0$, implies $\nabla p = \rho \mathbf{g}$. Choosing z to point upwards with $z = 0$ at the free surface, this leads for incompressible media to

$$p(z) = p_0 - \rho g z \qquad \text{or} \qquad \psi_m(z) = -\rho g z \; . \tag{3.34}$$

Non-Newtonian Fluids For non-Newtonian fluids, a number of empirical viscosity models have been proposed to replace (2.11). A popular one is the power-law $\Pi_{yx} = -\beta|\mathrm{d}_y v_x|^{n-1}\mathrm{d}_y v_x$ with coefficients β and n. One can show that this also leads to a power-law for the macroscopic velocity,

$$|\mathbf{j}_w|^{n-1}\mathbf{j}_w = -\frac{1}{\mu_{\text{eff}}}\mathsf{k}\nabla p \; , \tag{3.35}$$

where the coefficients n and μ_{eff} depend on the fluid [*Bird et al.* 1960].

Weak Inertia For flow that cannot be considered as slow, hence $\text{Re} \approx \text{Re}_{\text{crit}}$, the effect of inertia is not yet dominating but also cannot be neglected. Probably the first modification of Darcy's law for this regime was proposed by *Forchheimer* [1901]. For one-dimensional flow in a uniform and isotropic porous media, he suggested the empirical relation

$$j_w + \beta j_w^n = -K\partial_x p \; , \tag{3.36}$$

where β, n, and K are coefficients that depend on the shape of the pore space. Using a homogenization approach, *Mei and Auriault* [1991] found $n = 3$ and concluded that Darcy's law is thus also applicable for moderately high flow velocities.

Non-Stationary Flow While weak inertia and non-Newtonian behavior can be handled at least in an empirical way, the situation is more severe for non-stationary flow. If the external time scale that is given by the temporal change of the pressure at the boundary ∂V is comparable or even shorter than the internal time scale which is given by the dynamics inside V, we cannot expect a functional relation between $\mathbf{j}_w$ and ∇p anymore. Then, the very formulation of a flux law becomes impossible. This is generally not a problem for sufficiently small volumes and slowly changing external forcing. However, it becomes a dominating issue when the reference volume increases, as is the case when moving from the laboratory through the field to the regional scale, and if the forcing is fast as for instance when starting to pump a well or with the onset of a strong rainfall event after a dry period.

3.4
Material Properties

Combining the conservation of mass, for instance (3.25), and Darcy's empirical flux law (3.33) readily yields a generic formulation for the macroscopic dynamics of fluids in porous media. Using the matric potential ψ_m instead of pressure p and dropping subscripts this may be written as

$$\begin{aligned} \partial_t[\theta\rho] + \nabla\cdot[\rho\mathbf{j}] &= 0 \\ \mathbf{j} &= -\mathsf{K}[\nabla\psi_m - \rho\mathbf{g}] \,. \end{aligned} \tag{3.37}$$

Obviously, this formulation must be supplemented by descriptions of the material properties: the relation between θ and ψ_m, the hydraulic conductivity K, and possibly the density ρ in case the fluid is compressible.

Here, a remark is in order: The description of temporal changes of fluid content and divergence of flow at the macroscopic scale by (3.37) is based on the *local equilibrium hypothesis*, on the assumption that microscopic processes are fast on the time scale of the macroscopic processes and may thus be considered as being in equilibrium. Material properties, which reflect the action of microscopic processes at the macroscale, thus must not depend on the speed of macroscopic evolution. If experimental evidence shows that this hypothesis is not satisfied, then (3.37) is not valid.

3.4.1
Capacity

The pressure jump across an interface – the matric potential – is related to the mean curvature of the interface by (3.2). This curvature in turn is bound by the radius of the respective pore and by the contact angle. Hence, we expect for a porous medium a relation between the water-saturated volume fraction θ and the matric potential ψ_m and define the *hydraulic* or *soil water capacity* as

$$C(\psi_m) := \frac{\partial\theta}{\partial\psi_m} \,. \tag{3.38}$$

Traditionally, the integral of this function, the fluid fraction

$$\theta(\psi_m) = \theta_0 + \int_0^{\psi_m} C(\psi_m')\,\mathrm{d}\psi_m \tag{3.39}$$

as a function of the matric potential is more popular and graphical representations almost invariably show $\theta(\psi_m)$, not $C(\psi_m)$. Depending on context and discipline $\theta(\psi_m)$ is referred to as *soil water characteristic*, *pressure-saturation relation*, or *desaturation-imbibition curve*.

Bundle of Capillaries Consider a bundle of vertical capillaries in Figure 3.19. With their lower ends in a free wetting fluid, the height of the menisci is given by (3.6). Above the capillary fringe – given by the capillary rise in the largest pore –, the fluid fraction decreases monotonically with height. Consider a thin horizontal section with height Δz. According to (3.34), the matric potential changes by $\Delta\psi_m = -\rho g\Delta z$ across that height.

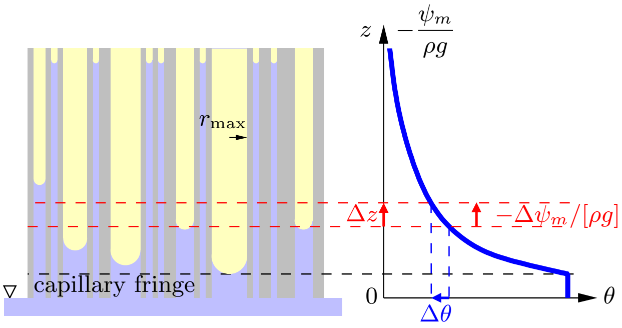


Figure 3.19. Hydraulic capacity of porous medium illustrated for a bundle of capillaries.

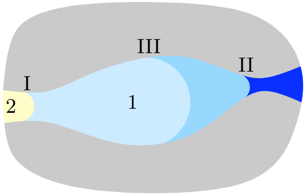


Figure 3.20. Hysteresis of hydraulic capacity in natural porous media during invasion of fluid 2 into pore initially filled with fluid 1, and vice versa.

Denoting the corresponding change of the fluid fraction by $\Delta\theta$, we define the hydraulic capacity C as the ratio $\Delta\theta/\Delta\psi_m$, in the limit $\Delta\psi_m \to 0$.

We thus already recognize the important fact that $\theta(\psi_m)$ can be read directly from the fluid fraction for a stationary situation above a free fluid table. This remains true also for natural porous media.

We notice in passing, that the distribution of radii of a bundle of capillaries can be calculated easily from measurements of $C(\psi_m)$. This cannot be extended to natural porous media, however.

Porous Medium While $C(\psi_m)$ and $\theta(\psi_m)$ are unique functions for a bundle of cylindrical capillaries, the situations becomes much more complicated for a general porous medium where the pore radius is almost never constant. To understand the impact of this, we consider a single pore that is wider in the middle (Figure 3.20). Starting from a pore that is initially saturated with fluid 1, pressure at one end shall decrease gradually. Once it falls below the *entry pressure* for fluid 2, the interface moves gradually into the pore up to the next pore throat, where the pore radius is minimal. This corresponds to meniscus I in Figure 3.20. Reducing the pressure any further empties the entire cavity and leads to meniscus II because the pore radius in the cavity is too large to sustain the interface whose radius is determined by the pressure jump across the interface. Reverting at this point and gradually increasing the pressure will not lead back to I, however. Instead, fluid 1 will invade the cavity until the pressure is increased such that the jump at the interface corresponds to the largest radius, menuscus III. Increasing it any further has fluid 1 fill the entire cavity and actually also the adjacent throat because the pore radii are smaller than the interfacial radius. Such discontinuous changes of the fluid content are referred to as *Haines jumps* [*Haines* 1930]. Understanding a single pore, we expect that for a porous medium both the

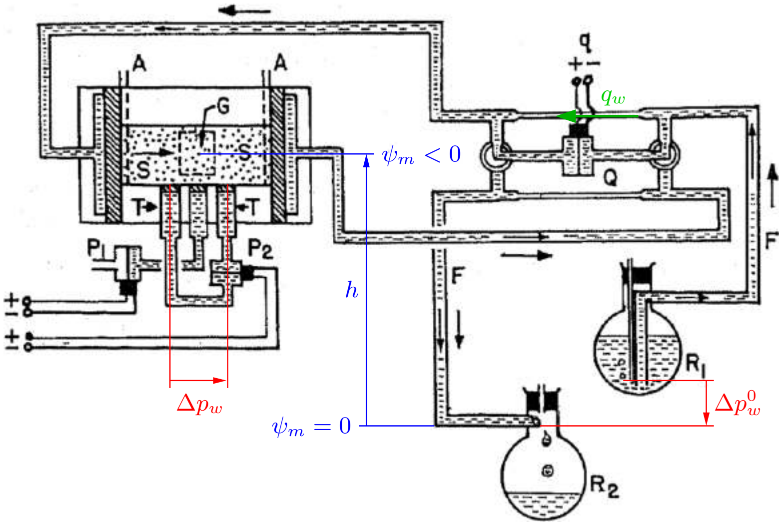


Figure 3.21. Experimental setup used by *Topp and Miller* [1966] for the direct measurement of soil hydraulic properties of sample S. Instrumentation includes tensiometers T with pressure transducers P_1 and P_2, γ-ray transmission with sample region G, capillary flow meter Q, and water reservoirs R_1 and R_2. Notice that reservoir R_1 is used as a Mariotte bottle that supplies water at a constant pressure independent of the fill level. Air inlets A ensure that air pressure in the sample is uniform and equal to ambient pressure.

hydraulic capacity $C(\psi_m)$ and also $\theta(\psi_m)$ are hysteretic. This was indeed already found in the first experiments reported by *Haines* [1930] and *Richards* [1931] and it has been studied in great detail ever after.

Topp and Miller [1966] did some of the early measurements of air-water systems in flow cells filled with various mixtures of glass beads (Figure 3.21). They used a horizontal flow cell in which the matric potential ψ_m could be adjusted in a wide range by varying the height h. Pressure p_a in the air phase was uniform and equal to that of the ambient air due to air inlets A. Water content θ was measured with γ-ray attenuation and $\psi_m = p_w - p_a$ in the sample was directly obtained from pressure transducer P_1. Some directly measured soil water characteristics $\theta(\psi_m)$ are shown in Figure 3.22.

We first focus on wetting or imbibition cycles. As a preparatory step to reach a well-defined initial state, the medium is first completely saturated with water and then drained by reducing the matric head from 0 to about -65 cm. This is the so-called *primary drainage branch*. The imbibition branch that starts from this end point does not reach a complete saturation anymore: the saturation only reaches about 0.85, even for $p = 0$. This results from *residual air*, i.e., from air which is completely enclosed by water and does not form a connected phase anymore. The next draining produces the *secondary drainage branch* which is now only followed up to a certain matric

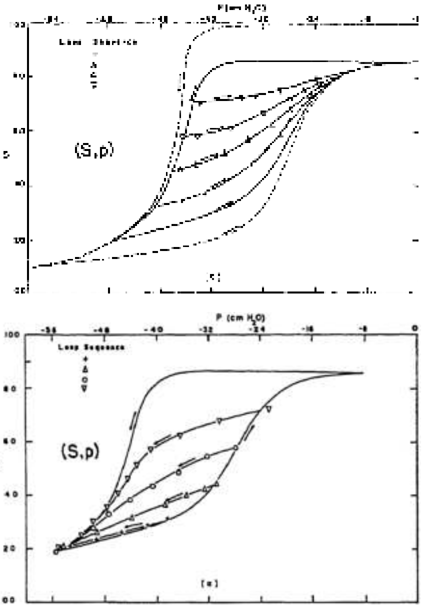


Figure 3.22.
Pressure-saturation relation measured by *Topp and Miller* [1966] with the apparatus shown in Figure 3.21. The sample consisted of packed glass beads with mean diameter 180 μm. **Top**: Wetting loops, with dashed line indicating primary drying. **Bottom**: Drying loops. The arrows indicate the direction of movement.
Notice that pressure p is given as equivalent height of a water column, i.e., as matric head. Saturation S equals the ratio θ/ϕ of water content and porosity. Compared to the representation in Figure 3.19, the curves are rotated by 90°. [*Figure combined from figures 2 and 3 of Topp and Miller [1966].*]

head before imbibition is initiated again. Various of these so-called *scanning loops* are shown.

Next, we consider the drying or desaturation cycles that start from different locations on the imbibition branch and approach the secondary drainage branch. Apparently, the wetting and drying loops are different and resemble in shape the respective main branches that also limit them.

What is the reason for the different shapes of the hysteresis loops? The initial preparation of the system was such that water filled the entire pore space. Through drainage and re-wetting a very complicated configuration of the water phase evolves. As long as both, the water- and the air-phase are connected, the curvature of their interface is the same everywhere and determined by the matric head. As is illustrated in Figure 3.20, this condition can be satisfied by many different configurations each of which is associated with a different water content. Moving the system through a succession of states along equilibrium paths, we thus cannot expect it to return to the same configuration by merely re-establishing the initial matric head. Also from Figure 3.20 we expect that the drainage branches reflect the pore volume associated with bottle-necks while, vice versa, imbibition branches reflect the cavities.

We mention in closing that the experiments could also be started from a completely air-saturated (dry) state. This would lead to the *primary imbibition branch* and to corresponding scanning branches. Such experiments are difficult, however, because the associated equilibration times are much longer. As will be shown below, this is due to the greatly reduced hydraulic conductivity at low water saturations.

Parameterizations Traditionally, the soil water characteristic $\theta_w(\psi_m)$ is parameterized instead of the soil water capacity $C_w(\psi_m)$. Its parameterization has a lengthy history. Over long stretches, it was driven by the desire to relate $\theta_w(\psi_m)$ to properties of the porous material, or at least to estimate it from more easily observable quantities. The prime quantity $\theta_w(\psi_m)$ was aimed at is the pore-size distribution that would be used for predicting other processes like chemical sorption or microbial activity. So far, this endeavor has not been successful because the pore-size distribution is only one part of the microscopic basis of $\theta_w(\psi_m)$. The other one is the connectivity between different pore-size classes, i.e., the topology of the pore space. Indeed, numerical simulations demonstrated that the two can be varied in a wide range while still leading to the same soil water characteristic [*Vogel* 2000]. While one can imagine to deduce both relations simultaneously from comprehensive measurements of $\theta_w(\psi_m)$ including various hysteresis loops, this is experimentally cumbersome and the direct imaging with x-ray microtomography will be preferred [*Okabe and Blunt* 2007; *Costanza-Robinson et al.* 2008].

Also the lesser goal, estimating $\theta_w(\psi_m)$ from more easily observable quantities, could only be reached in part. *Arya and Paris* [1981] estimated it from grain-size distributions, which is found to be useful for coarse-grained, unconsolidated media [*Fredlund et al.* 2002]. This method does not work for the more typical fine-textured, aggregated soils, however. For them, a statistical approach was followed with the premiss that soil hydraulic properties depend in a reasonably strong way on soil texture, density, content of organic matter, and possibly further soil bulk properties. Various regression relations, so-called *pedotransfer functions*, were then proposed [*Vereecken et al.* 1989; *Cornelis et al.* 2001; *Wösten et al.* 2001]. While several of them are useful, for instance for constructing some simulation scenarios, they hardly ever yield the accuracy required for representing flow, and possibly transport, at a particular site.

Given this state of affairs, we will follow a pragmatic line by accepting $\theta_w(\psi_m)$ as purely empirical macroscopic description which contains some indications on the underlying microscopic quantities – metric and topology of pore space –, but which does not allow to disentangle them quantitatively. We will in the following look into some common parameterizations of single hysteresis branches of $\theta_w(\psi_m)$. Current approaches for representing the full hysteresis use these same parameterizations and only adjust the parameters between upon a transition from one branch to another.

It is convenient to expressed the parameterizations in terms of water saturation and matric head. The *water saturation* is thereby defined as

$$\Theta := \frac{\theta - \theta_r}{\theta_s - \theta_r} \ , \tag{3.40}$$

where θ_r and θ_s are the residual and the saturated water content, respectively, and where the subscript w was dropped. Notice that these connotations of θ_r

and θ_s are strictly applicable only in the case of the primary drainage branch. In all other cases, they just represent the lower and upper end of the range of water contents. The matric head

$$h_m := \frac{\psi_m}{\rho_w g} = \frac{p_w - p_0}{\rho_w g} \tag{3.41}$$

was already defined by (3.19)–(3.20). Here, p_0 is the pressure in the air phase which is constant in the Richards regime. Recall $h_m < 0$ for bound water.

There exists a whole swath of parameterizations and we will only consider three of them: (i) the Brooks-Corey parameterization which is characterized by a sharp air-entry value and a correspondingly discontinuous soil water capacity $C_w(\psi_m)$, (ii) the van Genuchten parameterization which is very smooth throughout, most popular and available in all simulation codes, and (iii) the modified van Genuchten parameterization which allows a smooth transition between the previous two and which amends an issue of the original van Genuchten formulation for media with a wide pore-size distribution. Besides these parameterizations with fixed shapes, more flexible ones are gaining popularity [*Prunty and Casey* 2002; *Pedroso et al.* 2009].

Brooks-Corey Parameterization Introduced by *Brooks and Corey* [1966], this parameterization assumes a power-law distribution for the equivalent pore radii with a finite upper limit that corresponds to a maximum pore-size. This leads to a sharp air-entry value h_0 – the value required to empty the largest pore, thereby allowing air to enter the porous medium –, and to the formulation (Figure 3.23)

$$\Theta(h_m) = \begin{cases} [h_m/h_0]^{-\lambda} & ; h_m < h_0 \ , \\ 1 & ; h_m \geq h_0 \ , \end{cases} \tag{3.42}$$

where $h_0 < 0$ acts as a the scaling parameter. The exponent $\lambda > 0$ describes the distribution of the equivalent pore radii. For $\Theta < 1$, (3.42) may be inverted, yielding

$$h_m(\Theta) = h_0 \Theta^{-1/\lambda} \ . \tag{3.43}$$

Van Genuchten Parameterization While the Brooks-Corey parameterization appears attractive for describing a bundle of capillaries, it has two major drawbacks: (i) the soil water capacity $\mathrm{d}_h\theta$, which is a crucial parameter in (6.3), is discontinuous at $h_m = h_0$ and (ii) recalling the averaging inherent in the REV, one might question the sharp air-entry value. Both issues are resolved by the *van Genuchten* [1980] parameterization

$$\Theta(h_m) = \left[1 + [\alpha h_m]^n\right]^{-m} \tag{3.44}$$

with parameters $\alpha < 0$, $n > 1$, and $m > 0$ or, in the often used simplified form,

$$\Theta(h_m) = \left[1 + [\alpha h_m]^n\right]^{-1+1/n} \tag{3.45}$$

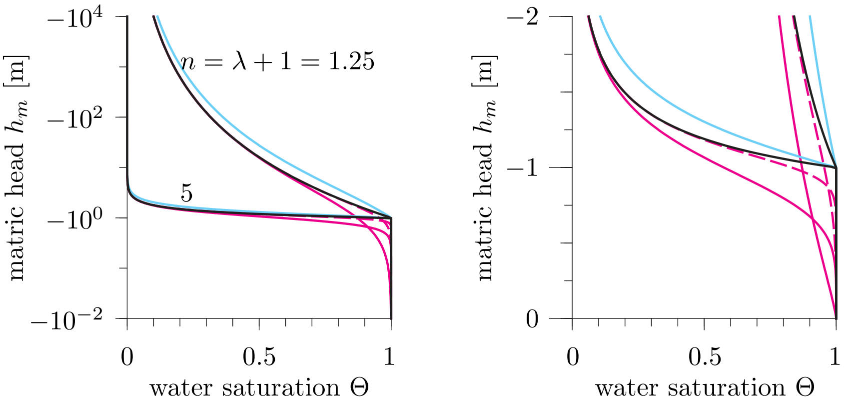


Figure 3.23. Brooks-Corey (black), van Genuchten (magenta), and modified van Genuchten (cyan) parameterizations of the soil water characteristic $\Theta(h_m)$ in semi-log (left) and linear (right) representations. The dashed magenta line corresponds to (3.44) and the solid one to the simplified version (3.45). Parameters are chosen such that curves approach each other for $h_m \to -\infty$. Specifically, $1/\alpha = h_0 = -1$ m for all models, $n = \lambda + 1$ for the simplified form (3.45) and $n = 3[\lambda + 1]$ with $mn = \lambda$ for the full form (3.44). For modified van Genuchten parameterization, $h_e = h_0 = -1$ m with the other parameters equal to the full form.

with $m = 1 - 1/n$ (Figure 3.23). We recognize α as scaling parameter for the matric head with $1/\alpha$ analogous to the air entry value. Notice, however, that air-entry is not sharp anymore and the very notion looses its meaning for small values of n. For $\alpha h_m \gg 1$, (3.44)–(3.45) approach a power function with exponent $-mn$ and $-[n-1]$, respectively. Comparison with (3.42) shows that the two parameterizations approach each other for $\alpha h_m \gg 1$ and for the parameter associations

$$\frac{1}{\alpha} = h_0 \qquad \text{and} \qquad \{mn = \lambda \quad \text{or} \quad n - 1 = \lambda\} \,. \tag{3.46}$$

Further looking into the parameters of (3.44) shows that n determines the sharpness of the air-entry and mn the distribution of equivalent pore radii (magenta curves in Figure 3.23). These two get linked together in the simplified formulation (3.45). Finally, inverting (3.44)–(3.45) yields

$$h_m(\Theta) = \alpha^{-1}\left[\Theta^{-1/m} - 1\right]^{1/n} \tag{3.47}$$

and

$$h_m(\Theta) = \alpha^{-1}\left[\Theta^{-n/[n-1]} - 1\right]^{1/n} \,, \tag{3.48}$$

respectively.

Modified Van Genuchten Parameterization An issue with the original van Genuchten parameterization is its shape near $h_m = 0$ and in particular the permissible range for n. Requiring the soil water capacity to be continuous at $h_m = 0$, which certainly is a physical prerequisite, *Luckner et al.* [1989] found $n > 1$. For $C_w(\psi_m)$ to be smooth at $h_m = 0$, they found $n > 2$ but also noted that many fine-textured soils were best described by $1 < n < 2$. As it turns out, the full van Genuchten parameterization (3.44) readily solves this apparent contradiction because it separates the parameterization of the air-entry region, parameter n, from that of the equivalent pore-size distribution at large, parameter m. This is illustrated by the dashed line in Figure 3.23.

Ippisch et al. [2006] chose a more radical solution and modified the parameterizations (3.44)–(3.45) such that a sharp air-entry occurs as it is found in the Brooks-Corey formulation. They introduce an additional parameter h_e, the air-entry head, cut $\Theta(h_m)$ at that point, and scale it linearly such that $\Theta(0) = 1$. This leads to

$$\Theta(h_m) = \begin{cases} \dfrac{\Theta_{\mathrm{vG}}(h_m)}{\Theta_{\mathrm{vG}}(h_e)} & ; h_m < h_e \ , \\ 1 & ; h_m \geq h_e \ , \end{cases} \tag{3.49}$$

where $\Theta_{\mathrm{vG}}(h_m)$ is one of (3.44)–(3.45) and again $h_m < 0$ (cyan curves in Figure 3.23). This modification obviously reintroduces the difficulty of the Brooks-Corey parameterization that the soil water capacity is discontinuous at h_e. A solution to this, equally applicable to the Brooks-Corey parameterization itself, is a heuristic regularization in a small environment of this point, i.e., a smoothing of the sharp cusp.

Comments We recall that parameterizations are only objective to the extent to which they refer to an appropriate REV. In particular, they are based on the assumption that all microscopic quantities are in equilibrium at the scale of the REV.

For simplicity, the parameterizations have been defined for the entire range of water contents. However, it is important to realize that they are *physically meaningful* only to the extent that the quantity they describe is well-defined. For the soil water characteristic considered here and for the hydraulic conductivity in the following Section, this typically implies that both the water and the air phase are continuous and that in addition the air phase is much more mobile than the water phase. This is the so-called *degenerate multiphase* or *Richards* regime that will be introduced in Section 6.1.

3.4.2
Conductivity

The hydraulic conductivity relates two vector quantities and is thus in general a second order tensor. While its directional nature has been addressed in a

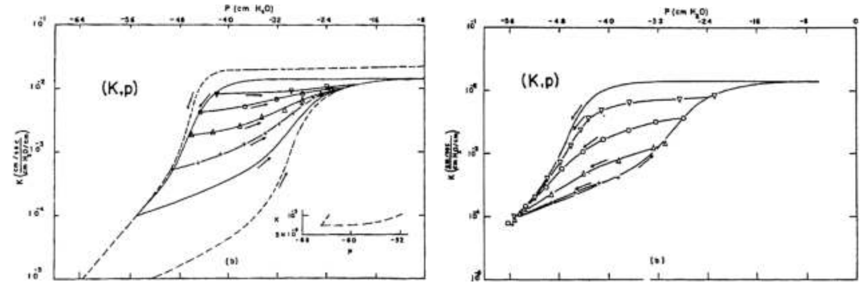


Figure 3.24. Pressure-conductivity relation corresponding to Figure 3.22. **Left**: Wetting Cycles. **Right**: Drying Cycles. Inset shows primary drainage and main imbibition curves. [*Figure combined from figures 2 and 3 of Topp and Miller [1966].*]

number of theoretical and numerical studies in the past, there is a paucity of experimental data on this aspect. Almost exclusively K is presumed to be isotropic and only the scalar K is measured, typically in a variant of the setup shown in Figure 3.21. There, one has two controls on the flow: (i) the pressure difference Δp_w^0 between the inlet and the outlet end which drives the water flow q_w and (ii) the height h of the sample above the reference plane which determines the matric potential ψ_m. This allows to independently vary ψ_m and its gradient $\partial_x\psi_m$. These two quantities are measured with the two pressure transducers P_1 and P_2. Relating $\partial_x\psi_m$ to the flow q_w, which is obtained from the pressure drop across a calibrated capillary, yields the hydraulic conductivity $K = -q_w/[A\partial_x\psi_m]$, where A is the cross-sectional area of the sample tube.

As we found previously, (3.30), the conductivity of a medium saturated with a single fluid is a conglomerate of the medium's permeability and of the fluid's dynamic viscosity. How does K then change in the presence of a second fluid and is it a unique function of matric potential ψ_m or of fluid fraction θ? We first convince ourselves that $K(\psi_m)$ is not unique by looking at Figure 3.20 and envisaging flow perpendicular to the drawing plane. Experimental evidence supports this (Figure 3.24).

Whether $K(\theta)$ is a unique function is difficult to answer intuitively. However, experimental evidence reveals that this is indeed the case to a fair degree of accuracy (Figure 3.25). Evidently, the conductivity varies over several orders of magnitude even for rather moderate variations of θ. There are three main causes for this: With decreasing saturation (i) the cross-sectional area decreases, (ii) the water phase is restricted to every smaller pores, and (iii) the microscopic potential gradient decreases because the path within the water phase between any two points becomes larger.

Parameterizations According to (2.38) and Buckingham's conjecture, the hydraulic conductivity $\mathsf{K}(\theta)$ is proportional to θ and to the mean cross-sectional area ℓ^2 of the flow channels, and inversely proportional to the dynamic viscosity. The constant of proportionality generally varies with θ

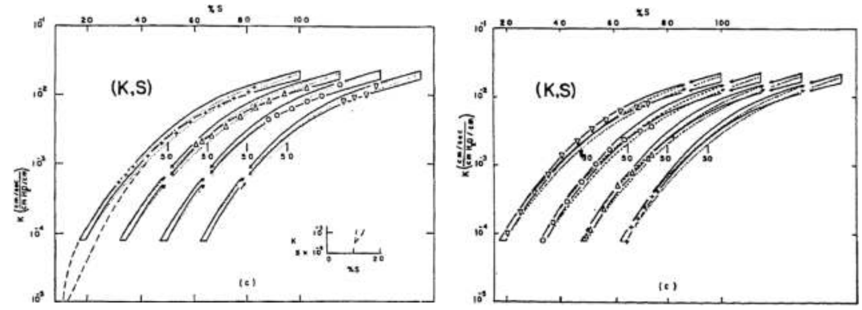


Figure 3.25. Saturation-conductivity relation corresponding to Figures 3.22–3.24. **Left**: Wetting Cycles. **Right**: Drying Cycles. Notice that the cycles practically overlay but are shifted relative to each other in order to separate them. [*Figure combined from figures 2 and 3 of Topp and Miller [1966].*]

and depends on many intricate properties, among them the distribution of ℓ^2, the tortuosity of the flow channels, and their topology. Information on these properties is available, in implicit form, from $\Theta(h_m)$.

Historically, a succession of ever more refined models of the pore space were studied. They all assumed the medium to be isotropic and deduced a single function $K(\Theta)$ based on $h_m(\Theta)$ and some additional parameters. The starting premise is that, with (3.6),

$$\frac{\mathrm{d}h_m}{\mathrm{d}\Theta} = -\frac{2\sigma_{wa}}{\rho_w g R^2}\frac{\mathrm{d}R}{\mathrm{d}\Theta} \tag{3.50}$$

yields the distribution of pore radii R of the water-filled fraction of the pore space and that the resulting conductivity function may be calculated from Hagen-Poiseuille's solution (2.35). *Childs and Collis-George* [1950] considered randomly connected stacks of capillary bundles, assumed the conductivity of a single path to be determined by its smallest radius and neglected the impact of bifurcations. *Mualem and Dagan* [1978] showed that this model, together with others that make additional assumptions on the tortuosity, may be cast into the form

$$K(\Theta) = K_0\Theta^a \frac{\int_0^\Theta [\Theta - \vartheta] h_m(\vartheta)^{-2}\,\mathrm{d}\vartheta}{\int_0^1 [\Theta - \vartheta] h_m(\vartheta)^{-2}\,\mathrm{d}\vartheta}\ , \tag{3.51}$$

where the *saturated conductivity* K_0 and a are parameters, with a often associated with the tortuosity of the pore space.

Randomly choosing the larger or the smaller radius of a composite pore as its effective radius yields the parameterization

$$K(\Theta) = K_0\Theta^a \frac{\int_0^\Theta h_m(\vartheta)^{-2}\,\mathrm{d}\vartheta}{\int_0^1 h_m(\vartheta)^{-2}\,\mathrm{d}\vartheta}\ , \tag{3.52}$$

which was originally proposed by *Burdine* [1953]. Here, Θ^a again accounts for the tortuosity. A further alternative for the choice of an effective radius was put forth by *Mualem* [1976] who assumed geometrically similar pores which leads to the geometric mean of the individual radii. He then obtained the parameterization

$$K(\Theta) = K_0\Theta^a \left[\frac{\int_0^\Theta h_m(\vartheta)^{-1}\,\mathrm{d}\vartheta}{\int_0^1 h_m(\vartheta)^{-1}\,\mathrm{d}\vartheta}\right]^2 . \tag{3.53}$$

Current practice prefers this last model and it is also implemented in most numerical Richards solvers. It can be combined with any of the $\Theta(h_m)$ models to arrive at a coherent parameterization of the soil hydraulic properties. Indeed, some parameterizations of $\Theta(h_m)$ have been restricted purely on the grounds that an analytic solution of the integrals in (3.53) is feasible. This is most prominently the case with the choice $m = 1 - 1/n$ for the simplified van Genuchten parameterization (3.45) [*van Genuchten* 1980]. With a numerical evaluation of (3.53), this is no more a limitation, however.

Apparently, there is considerable latitude in the formulations (3.51)–(3.53). On the one hand, it stems from the definition of an effective pore radius. On the other hand, we recall that experimental evidence shows only a minor hysteresis for $K(\Theta)$, but a major one for $h_m(\Theta)$. There is no provision in (3.51)–(3.53) to account for this, even though different hysteresis branches lead to quite different functions $h_m(\Theta)$. So far, this issue has not been explored to any depth.

We will take a pragmatic approach to this gray area and in the following employ the various parameterizations for studying qualitative aspects only. To this end, it suffices to consider the Mualem-Brooks-Corey- and the Mualem-van Genuchten-parameterization.

Mualem-Brooks-Corey Inserting (3.43) into (3.53) yields

$$\begin{aligned} K(\Theta) &= K_0\Theta^{a+2+2/\lambda} \\ K(h_m) &= \begin{cases} K_0[h_m/h_0]^{-2-\lambda[a+2]} & ; h_m < h_0 \ , \\ K_0 & ; h_m \geq h_0 \ . \end{cases} \end{aligned} \tag{3.54}$$

Notice that $K(h_m)$ is a straight line for $h_m < h_0$ in a double logarithmic representation,

$$\log\Bigl(\frac{K(h_m)}{K_0}\Bigr) = \bigl[-2 - \lambda[a+2]\bigr] \log\Bigl(\frac{h_m}{h_0}\Bigr) . \tag{3.55}$$

Mualem-van Genuchten Following *van Genuchten* [1980] and inserting (3.48) into (3.53) yields

$$\begin{aligned} K(\Theta) &= K_0\Theta^a\Bigl[1 - \bigl[1 - \Theta^{n/[n-1]}\bigr]^{1-1/n}\Bigr]^2 \\ K(h_m) &= K_0\bigl[1 + [\alpha h_m]^n\bigr]^{-a[1-1/n]} \\ &\qquad \Bigl[1 - [\alpha h_m]^{n-1}\bigl[1 + [\alpha h_m]^n\bigr]^{-1+1/n}\Bigr]^2 . \end{aligned} \tag{3.56}$$

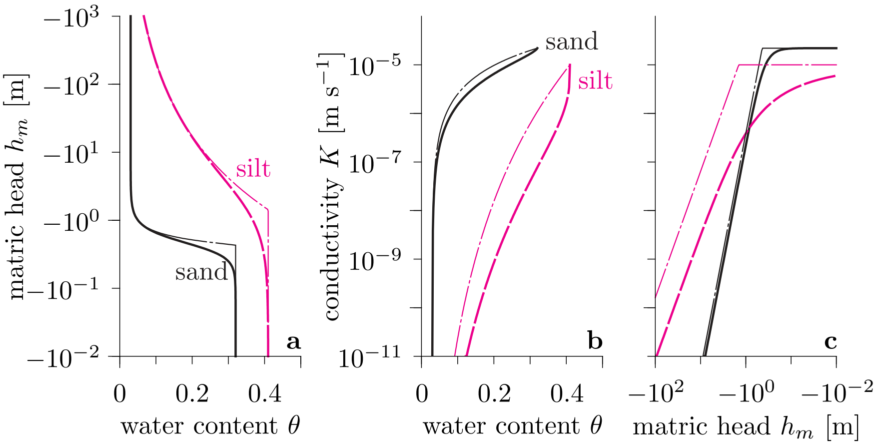


Figure 3.26. Soil hydraulic properties for two different soil textures in the Mualem-van Genuchten (thick lines) and in the Mualem-Brooks-Corey (thin dash-dotted lines) parameterization: soil water characteristic function $\theta(h_m)$ **(a)** and the hydraulic conductivity function $K(\theta)$ **(b)** and $K(h_m)$ **(c)**. The parameters used are adapted from *van Dam et al.* [1992] and are given in Table 3.1.

For $\alpha h_m \gg 1$, $K(h_m)$ again approaches a straight line in a double logarithmic representation,

$$\log\left(\frac{K(h_m)}{K_0}\right) \approx 2\log\left(1-\frac{1}{n}\right) + \big[a - n[a+2]\big]\log(\alpha h_m) \ . \tag{3.57}$$

Example: Typical Soil Hydraulic Properties Soils differ greatly in their composition of primary particles – with sizes ranging from a few millimeters for coarse sand to less than a micrometer for fine clay – and this is reflected in their hydraulic properties. In Table 3.1, values for the Mualem-van Genuchten parameterization of two typical soil materials reported by *van Dam et al.* [1992] are listed. The corresponding functions are shown in Figure 3.26 together with the Mualem-Brooks-Corey parameterization obtained with (3.46).

Table 3.1. Mualem-van Genuchten parameters for hydraulic properties $\theta(h_m)$, $K(\theta)$, and $K(h_m)$ shown in Figure 3.26. The parameters have been adapted from [*van Dam et al.* 1992] who fitted them to measured data. The last column is the slope $\kappa = a - n[a+2]$ of $K(h_m)$ in a double logarithmic plot for $\alpha h_m \gg 1$ as given by (3.57). It is identical to the slope of $K(h_m)$ in the Mualem-Brooks-Corey model (3.55).

	θ_r	θ_s	α [m^{-1}]	n	a	K_0 [m s^{-1}]	κ
sand	0.03	0.32	−2.3	4.17	−1.1	$2.2{\cdot}10^{-5}$	−4.9
silt	0.01	0.41	−0.7	1.30	0.0	$1.0{\cdot}10^{-5}$	−2.6

The soil water characteristic function reveals that, as expected, in the silt, the fine-textured medium, the air-entry value $1/\alpha$ is higher than in sand. It also shows a much broader pore size distribution, i.e., θ changes over a range of $h_m \propto 1/r$ that is several orders of magnitude larger. For sand, θ ranges between θ_r and θ_s for h_m roughly in the interval $-2 \ldots -0.1$ m. With (3.6), this corresponds to structures of the pore space in the range $7 \ldots 150$ μm. In contrast, about 50% of the pore volume of the fine-textured medium is connected through pores with radii smaller than 0.7 μm.

The saturated conductivity K_0 of the coarse-textured sand is higher than that of the silt because (i) $K \propto \theta\ell^2$, (ii) θ at saturation is roughly the same for the two media, and (iii) the pore radii of the sand are significantly larger. As the media become unsaturated, the conductivity representations $K(\theta)$ and $K(h_m)$ reflect the simultaneous influence of the cross-sectional area on the one hand and of the pore radii on the other.

In the sand, $K(\theta)$ decreases rather gradually, by about an order of magnitude down to $\theta \approx 0.1$. As can be deduced from the soil water characteristic, θ and ℓ contribute roughly equally to this decrease. For $\theta < 0.1$, $K(\theta)$ drops dramatically because water is then only conducted by rapidly thinning films. The situation is quite different for the silt with its much broader pore size distribution. With the simplified van Genuchten parameterization (3.45), $K(\theta)$ drops by an order of magnitude with θ only decreasing slightly from saturation. This reflects the rapid decrease of the mean pore radius of the water-filled space with just slightly decreasing θ as it is apparent from the soil water characteristic function. *Ippisch et al.* [2006] pointed out that such a rapid decrease of $K(\theta)$ near saturation is not expected in real soils, since such fine-textured media would not contain any pores of the required size at all. They identified this as an artifact of the simplified van Genuchten parameterization for parameters $n < 2$. This issue can be resolved by using a form that allows the suppression of arbitrarily large pores. This is indeed achieved by all the other parameterization presented above.

Returning to the general behavior of $K(\theta)$ in the differently textured materials, notice that for a given value of θ, the conductivity of the silt is always smaller than that of the sand. *Vice versa*, for a given conductivity the water content of the silt is higher.

The behavior of $K(h_m)$ is very different from that of $K(\theta)$ since now the conductivity is considered as a function of the largest radius of the water-filled pore space. Hence, we essentially look at the cross-sectional area for a given radius. For the sand, θ decreases much more rapidly with decreasing radius than for the silt. Consequently, the same is true for the conductivity. We thus recognize that, in contrast to $K(\theta)$, the $K(h_m)$ curves of two differently textured porous media generally cross each other. This will turn out to be a crucial aspect of water flow in the vadose zone.

Concluding Remark Notice that the two parameterizations of the conductivity are markedly different mainly near saturation because of their very different distribution of the largest pores which are also most effective in conducting water. This difference is then perpetuated also into the strongly unsaturated range since, as (3.53) shows, the contribution from a particular class of pore-sizes is scaled by the effect of the entire ensemble of pores. Recall, however, that the underlying assumption of a Richards regime does not permit us to approach

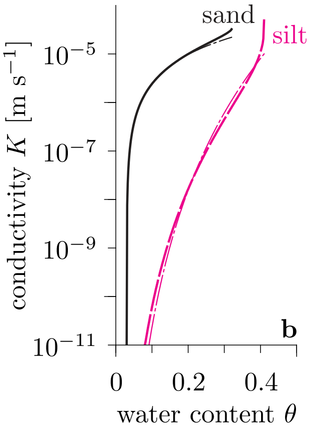


Figure 3.27.
Hydraulic conductivity functions $K(\theta)$ as shown in Figure 3.26 but with the Mualem-van Genuchten parameterization multiplied with a factor such that the two parameterization yield the same value at $\Theta = 0.5$.

the saturated range with this type of material properties. As a consequence, K_0 is a useful parameter for specifying a parameterization. Other than that it is of little practical significance for unsaturated media. In particular, measuring the saturated conductivity as it would be done for a groundwater study, e.g., with a falling head permeameter, and using this value to constrain estimates of the unsaturated conductivity $K(\theta)$ is a vain attempt which in general leads vast overestimations. Conversely, extrapolating some measurements of $K(\theta)$ to a state of near-saturation in order to estimate, for instance, the time to flooding during an intense storm, is equally useless and will lead to a severe underestimation. Reasons for both are (i) the limited applicability of Richards equation as mentioned above and (ii) the ubiquitous macrostructures in soils, e.g., cracks and wormholes, whose geometry is hardly related to the smaller-scale texture represented by the conductivity function, which are never active in the Richards regime, but can conduct very large fluxes once activated by flooding.

As a consequence, parameterizations of K that are obtained from or used for real situations are never related through K_0 but through $K(\theta_0)$, where $\theta_0 < \phi$ is a water content at which the Richards approximation is well-valid and large-scale structures of the pore space are negligible. Figure 3.27 shows $K(\theta)$ from Figure 3.26b with the Mualem-van Genuchten parameterization scaled accordingly.

Comment on Anisotropic Media Most experiments for measuring the hydraulic conductivity impose a constant gradient along the axis of some cylindrical sample and measure the resulting fluid flow. Unless the porous medium is indeed isotropic or at least one of its main axes is aligned with the cylinder axis, such a measurement does not provide a correct value for the conductivity component along the axis. The reason for this is that in the general case the anisotropy of K will lead to flow components orthogonal to the sample walls. Since this flow is impeded, a corresponding gradient of the hydraulic potential will be generated which in turn modifies the flow along the axis, again through the anisotropy of K. Denoting the axis of the sample by z and the two orthogonal directions by x and y, a quick calculation shows that the measurement yields

$$K'_{zz} = K_{zz} - \frac{K_{xz}^2 K_{yy} - 2K_{xy}K_{xz}K_{yz} + K_{xx}K_{yz}^2}{K_{xx}K_{yy} - K_{xy}^2} \; , \tag{3.58}$$

where K_{ij} is the true ij-component of K. Clearly, the measured value K'_{zz} only equals the true one if all off-diagonal elements vanish.

3.4.3
Flux Law

The empirical flux law for a single fluid was introduced in Section 3.3.5 as Darcy's law (3.30). The situation becomes much more difficult when more than one fluid is present. Here, *Buckingham's conjecture* is typically invoked which states that Darcy's law remains valid and the only modification required is to write K as a function of θ [*Buckingham* 1907]. Hence, for fluid i,

$$\mathbf{j}_i = -\mathsf{K}_i(\theta_i)\nabla\psi_i \; , \tag{3.59}$$

which is called the *Buckingham-Darcy law* in soil physics and hydrology and the *generalized Kozeny-Carman law* in petroleum industry [*Kozeny* 1927].

This law is remarkable for two reasons: (i) It states that the volume fraction of fluid i is the sole factor that modifies the conductivity. This is corroborated by experiments. (ii) It states that the flux in fluid i only depends on the hydraulic gradient in the same fluid. From Onsager's theorem, we might instead expect a flux law of the form

$$\mathbf{j}_i = -\sum_l \mathsf{K}_{il}(\theta_i, \theta_l)\nabla\psi_l \; , \tag{3.60}$$

where summation is over all phases l. This issue has not yet been explored, neither theoretically nor experimentally.

3.4.4
Compressibility

Porous media are generally compressible, many of them quite strongly as for instance most soils. In the following, we only consider weakly compressible media and think of water-saturated sandstone as a generic example. The compressibility of these media – water and porous sandstone – are so small that greatly simplifying assumptions are justifiable. In particular, we assume that the configuration of grains does not change with pressure p which implies that the load on the grain contacts is large compared to the pressure change.

Then, θ and ρ are simple functions of p and the storage term of (3.37) may be written as

$$\begin{aligned}\partial_t[\theta(p)\rho(p)] &= \left[\theta\frac{\mathrm{d}\rho}{\mathrm{d}p}+\rho\frac{\mathrm{d}\theta}{\mathrm{d}p}\right]\partial_t p = \rho\underbrace{\left[\theta\kappa_p^{\text{water}}+[1-\theta]\kappa_p^{\text{matrix}}\right]}_{\text{storage coefficient } S}\partial_t p\\ &= \rho S\partial_t p\ , \qquad (3.61)\end{aligned}$$

where the compressibility κ_p has been defined in (2.10) and S is the *volumetric storage coefficient*. Since the compressibilities are constant to a very good approximation, S is also constant.

Calculating the effect of compressibility on the storage term is straightforward. This is not the case for the flux term, however, since a large compressibility of the matrix constituents leads to a qualitative change of the pore space geometry. However, for the case of weakly compressible media like sandstones and sandy soils, simplifying approximations are permissible which make the problem tractable. Since we will find below, that the corresponding additional terms are negligible, the somewhat lengthy derivation is only given as the following parenthetical comment.

Detail: Conductivity of Weakly Compressible Media Write the flux term as $\nabla\cdot[\rho\mathbf{j}] = -\nabla\cdot\left[\rho K[\nabla p-\rho\mathbf{g}]\right]$, where p replaces ψ_m in (3.37) – we consider water-saturated media – and, for simplicity, the medium is isotropic. Then

$$\nabla\cdot\left[\rho K[\nabla p-\rho\mathbf{g}]\right] = \rho K\left[\nabla\cdot[\nabla p-\rho\mathbf{g}]+\frac{\nabla[\rho K]}{\rho K}\cdot[\nabla p-\rho\mathbf{g}]\right] \qquad (3.62)$$

with

$$\frac{\nabla[\rho K]}{\rho K} = \left[\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}p}+\frac{1}{K}\frac{\mathrm{d}K}{\mathrm{d}p}\right]\nabla p = \left[\kappa_p^{\text{water}}+\frac{1}{K}\frac{\mathrm{d}K}{\mathrm{d}p}\right]\nabla p\ . \qquad (3.63)$$

For the relative change of K with pressure we find, with $K=k/\mu$ (3.32),

$$\frac{1}{K}\frac{\mathrm{d}K}{\mathrm{d}p} = \frac{1}{\mu^{-1}}\frac{\mathrm{d}\mu^{-1}}{\mathrm{d}p}+\frac{1}{k}\frac{\mathrm{d}k}{\mathrm{d}p}\ . \qquad (3.64)$$

For liquids, μ is practically independent of p. Hence, it suffices to calculate the second term. We invoke the following deformation model which is only justifiable for weakly compressible media: grains deform isotropically except at their contacts where only the area shrinks (Figure 3.28). Then, the water phase remains geometrically similar with changing pressure and, since k is proportional to the square of the mean pore diameter ℓ, we obtain

$$\frac{1}{k}\frac{\mathrm{d}k}{\mathrm{d}p} = -\frac{2}{\ell}\frac{\mathrm{d}\ell}{\mathrm{d}p} = \frac{2}{3}\kappa_p^{\text{matrix}}\ . \qquad (3.65)$$

Here, the sign for the first equality indicates that the permeability increases with decreasing grain diameter, and the second equality follows from $V\propto\ell^3$ together with $\kappa_p=-\frac{1}{V}\frac{\mathrm{d}V}{\mathrm{d}p}$. Inserting everything into (3.63) yields

$$\frac{\nabla[\rho K]}{\rho K} = \left[\kappa_p^{\text{water}}+\frac{2}{3}\kappa_p^{\text{matrix}}\right]\nabla p\ . \qquad (3.66)$$

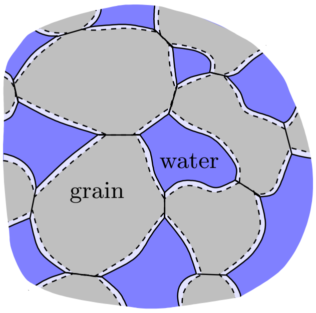


Figure 3.28.
In a simple model of weakly compressible media, the grains shrink isotropically with increasing pressure (dashed lines) except at the contact regions between grains where only the contact area decreases. Notice that shrinking is grossly exaggerated in this sketch and also porosity is rendered too large.

With this, a quick analysis of scales shows that for weakly compressible media as we consider them here, the second term in parenthesis of (3.62) is many orders of magnitude smaller than the first one. To a very good approximation, we may thus write for the flux term

$$\nabla \cdot [\rho \mathbf{j}] = -\nabla \cdot \big[\rho K [\nabla p - \rho \mathbf{g}]\big] = -\rho K \nabla \cdot [\nabla p - \rho \mathbf{g}] \;, \tag{3.67}$$

where ρ and K are now constant.

Example: Sandstone Aquifer Consider a typical formation with porosity $\phi = 0.1$ and compressibility $\kappa_p^{\text{matrix}} \approx 10^{-10}$ Pa^{-1}. The compressibility of sand grains is somewhat lower, about $3 \cdot 10^{-11}$ Pa^{-1}, with the higher value of the matrix reflecting the reduced contact area between the grains.

Storage Coefficient With $5 \cdot 10^{-10}$ Pa^{-1} for the compressibility of water, the storage coefficient becomes $S \approx 1.4 \cdot 10^{-10}$ Pa^{-1}. Reducing the pressure in this formation by 10^5 Pa thus releases some 14 grams of water per cubic meter of aquifer, or some 140 ppm of the stored mass.

Conductivity With (3.65) and assuming μ as constant, the relative change of K is about 10^{-9} Pa^{-1}. To set this in proportion, we employ the hydrostatic pressure gradient $\mathrm{d}_z p = \rho g$ and find a relative increase of K with depth due to the compressibility of water and matrix of some 10^{-5} m^{-1}, which is quite negligible, indeed.

Notice that the treatment of the influence of compressibility on the storage term $\partial_t[\rho\theta]$ was quite different from that on the flux term: While both are exceedingly small – for the example of the sandstone $S \approx 1.4 \cdot 10^{-10}$ Pa^{-1}, $\mathrm{d}_p K/K \approx 10^{-9}$ Pa^{-1}, and $\mathrm{d}_p \rho/\rho \approx 5 \cdot 10^{-10}$ Pa^{-1} – the storage coefficient was retained in the formulation while the flux was approximated by its incompressible part. The justification, and necessity, for this is that with $S = 0$ the character of the dynamics changes qualitatively in that flow is always stationary, hence a pressure fluctuation would transverse the flow domain with infinite velocity. We will elaborate on this in Section 5.1.

Exercises

3.1 Capillary Rise Deduce (3.6) by (i) minimization of energy and (ii) balancing the pressure jump across the meniscus with the hydrostatic pressure in the liquid column.

3.2 Closed Capillary Consider a cylindrical capillary with length L and circular cross-section with radius R. Let it be closed at one end. How high does water rise if the material is perfectly water-wet and if air can be described as an ideal gas? Does this height depend on temperature?

3.3 Corrugated Capillary Consider the situation depicted in Figure 3.14 and discuss the change of the equilibrium points D and I upon change of the water table.

3.4† Dynamics of Capillary Rise Consider a cylindrical capillary with circular cross-section and radius R that is open at both ends and much longer than the maximum capillary rise of water. Envisage the capillary to be hold vertically and, at time $t = 0$, its lower end to being placed such that it just penetrates the surface of a flat water table. A meniscus will then rise and eventually reach height h above the water table. For a perfectly wettable capillary, h is given by (3.6).

1. Qualitatively discuss the movement of the water within and outside of the capillary. Short of actually calculating the dynamics notice that it will be governed by a certain time scale. Explain why this is the case and express it (qualitatively) as a function of the system variables R, σ,. . . .
2. Making reasonable assumptions deduce a differential equation for the movement of the meniscus and solve it. [The second part is greatly facilitated by using *Mathematica* or similar.]
3. Discuss the impact of (i) a non-zero wetting angle and (ii) a modulation of diameter. [Do the second part only qualitatively. The problem is best approached by considering the pressure at the lower end of the capillary, within the capillary on the one hand and far away but at the same depth beneath the water table on the other.]

3.5 Water Drop on a Hair Qualitatively describe the shape of a water droplet on a water-wet hair, i.e., on a thin cylinder.

3.6† Air Bubble Pouring water from some height into a glas, bubbles form, migrate to the surface and remain there for quite some time before they eventually blopp. Explain the life cycle of such a bubble and in particular its stability at the water-air interface!

3.7† A Special Tea Treat yourself to a black tea with about 1/3 of milk. Add some cinnamon and pepper for good taste and a small piece of butter for the Tibetan touch. Make sure that the cup is not completely filled initially. After the butter has melted, pour in more hot water from some height such that bubbles form. Observe the form and the movement of the bubbles and of the butter drops. Explain!

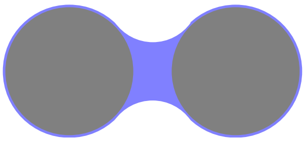


Figure 3.29.
Sketch for Exercise 3.8. Two spherical, water-wet grains of sand are fixed in space with some water extended between them. Since the grains are water-wet, they are covered by a film of water.

3.8 Two Grains with Water Consider Figure 3.29 and assume that the two grains are fixed in space. Discuss the shape of the air-water interface and explain if it is consistent with the Young-Laplace equation (3.2)? What force acts on one of the grains and what movement will it perform when it is free to move?

3.9 Cohesion of Sand Walking along a sandy beach, you notice that at a certain distance from the water, there extends a strip where you can walk easily without sinking in. Nearer to the water as well as farther away from it, this is not the case anymore and walking is difficult. Explain! [As an aside, the famous Daytona beach race in Florida takes much of its fascination from this effect. Obviously, leaving the optimal track with a fast car has spectacular consequences.]

3.10† Spilled Water Take a flat glass surface in your kitchen, wipe it clean and dry, and pour some water on it. Describe and explain the evolution and the final state of the spilled water. Then take a dry sponge cloth and lay it on the glass such that it covers a small part of the spilled water. Again, describe and explain the evolution and the final state of the spilled water.

You may want to play along the following lines:

1. Modify the glass' surfaces properties by cleaning it with a detergent, wiping it with some oily cloth, or painting some oily pattern on a previously detergent-cleaned surface. Watch the contact line move across the boundary between different surface properties as you continue to add water.
2. Add a second water spill and continue adding water until the two distributions touch. Be attentive to the moment of contact.
3. Continuously add the water as a fine spray, to the horizontal glass surface and also to the kitchen window. (The latter leads to a new field, self-organized criticality, described for instance by *Jensen* [1998].)

3.11 Divergence Show that divergence and integration may be interchanged in (3.21).

3.12 Units of Conductivity Darcy's law may be written in terms of the hydraulic potential, $\mathbf{j}_w = -\mathsf{K}\nabla\psi_w$, or in terms of the hydraulic head, $\mathbf{j}_w = -\mathsf{K}\nabla h_w$. For simplicity, the same symbol is often used for the hydraulic conductivity K. What are its units in the two formulations?

3.13 Conductivity Consider a uniform and isotropic porous medium with permeability $k = 10^{-10}$ m^2. (i) What is its hydraulic conductivity at 10°C and at 20°C? (ii) Let the medium be completely dry and consider air flow. Under what conditions can this be described by a conductivity in analogy to the hydraulic one? What is its value at 10°C? (iii) If the same medium were scaled up by a factor of 2, i.e., all lengths are stretched by a factor of 2, how do the answers to the previous questions change?

3.14† **"Conductivity Function" of Capillary** Consider a vertical, perfectly water-wet capillary with radius R. Assume a constant film flow in analogy to Exercise 2.3 and calculate the relation between film thickness d and water flux $j_w := q/[\pi R^2]$, where q is the flow.

Film flow in a capillary may be used as the most simple model for water flow in an unsaturated medium. Calculate and plot the relation between saturation Θ and conductivity K. Discuss.

3.15† **Hydraulic Properties of a Bundle of Capillaries** Consider a bundle of parallel capillaries with circular cross-sections. Assume the soil water characteristic $\theta(\psi_m)$ for this bundle to be given by the Brooks-Corey parameterization (3.42). Calculate (i) $n(r)$ $[\mathrm{m}^{-2}]$, the number of capillaries per unit cross-sectional area whose radius is smaller than r, and (ii) the hydraulic conductivity function $K(\theta)$. Discuss the applicability of this result to porous media.

3.16 REV and Fractal Media Sketch and discuss the porosity of the following media as a function of characteristic length ℓ of an isotropic averaging volume: (i) densest packing of uniform spheres with radius r, (ii) densest packing of two sorts of uniform spheres, one with radius r_1, the other one with radius r_2, where $r_1 \ll r_2$, (iii) a quasi-fractal distribution of mass with lower cutoff ℓ_1 and upper cutoff ℓ_2.

3.17† **Heating by Water Flow** Consider a vertical, 1 m long column uniformly filled with quartz sand. The column's porosity is 0.3 and the hydraulic conductivity $K_w = 10^{-7}\ \mathrm{m\,s^{-1}}$. It is water-saturated and the water table is held constant at the column's upper end. At the lower end, water flows out freely. The column is thermally isolated and initially at $T = 10^\circ\mathrm{C}$ as is the inflowing water. Assume that the water equilibrates thermally with the column as it flows through. Estimate the rate of temperature increase of the outflowing water.

3.18† **Kinetic vs Potential Energy** Calculate the kinetic energy of the flowing water in Exercise 3.17. To what height could the water be lifted with this energy? Repeat the calculation for the case of a storm flow event which leads to an infiltration of $20\ \mathrm{mm\,h^{-1}}$.

3.19 Hysteresis Consider the hysteretic soil water characteristic shown by Figure 3.22 on page 60. Take an arbitrary point on the main drainage branch and notice that for the scanning imbibition branch starting from there, the soil water capacity $C_w = \partial\theta/\partial\psi_m$ is much smaller than on the main drainage branch at the same point. Explain!

4
Solutes in Porous Media

A *solution* is a homogeneous mixture of two or more substances which form a single phase. While there exist solid solutions, they are not of interest here and we only consider fluid solutions, gaseous or liquid ones. Typically, the mass of the *solute* in a solution is a small fraction of the mass of the *solvent*. A solution is a fluid, solute and solvent are its components. Referring to *solution* instead of *fluid* indicates a focus that is more on chemical aspects than on physical ones.

Solute *transport* consists of two aspects: movement with the fluid and within it. The former originates from the fluid's flow which in turn is determined by the fluid's physical properties, by the geometry and physico-chemical properties of the bounding solid, and, in multiphase systems, by the properties of the further fluids. Movement within the fluid results from the thermal motion which leads to molecular diffusion. The two aspects of transport are referred to as *convection* and *diffusion*, respectively. While such a separation is straightforward in simple flow domains like capillaries, it is neither feasible nor desirable in larger porous formations where the flow field is typical exceedingly complicated. The effect of small-scale spatial variability of the flow velocity on large-scale solute transport is then only described statistically. The distinction between the resulting so-called *dispersion* and the explicitly represented convection, which may still vary at larger scales, is reminiscent of, and closely related to, the distinction between structure and texture of a porous medium as discussed in Section 3.3.2. Determining an optimal scale for the separation between convection and dispersion, and quantifying the latter, is the major challenge in studying the physics of solute transport.

4.1
Transport at the Pore-Scale

We consider a porous medium that is completely saturated with water and look into transport at increasingly larger scales. Starting at molecular diffusion in a pure fluid, we step up to transport with the laminar flow in

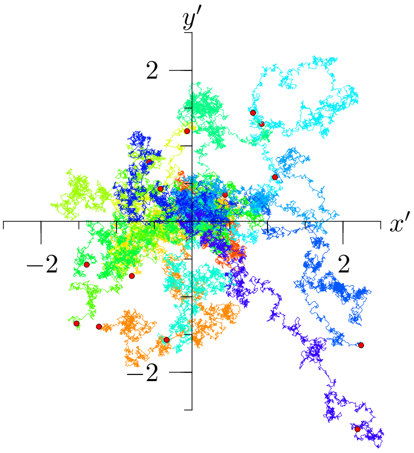


Figure 4.1.
Two-dimensional Brownian motion of 15 particles which all started at the origin at time $t = 0$. The axes are scaled with the standard deviation of the asymptotic probability density function, which is obtained from (4.2). Hence, $x' = x/\sqrt{2D_m t}$ and y' in analogy. The locations of the particles at the end of the path, after 3,000 random steps, are marked by a red dot.

a capillary, continue with mixing in pore junctions, and finally consider the experimental evidence for dispersion in porous media and formations.

As a preliminary, notice that there is a characteristic difference between these processes. Molecular diffusion is always active, with spreading essentially a function of time, and inherently isotropic, possibly restricted by phase boundaries. In contrast, the other processes are a direct consequence of fluid flow in the porous medium, with spreading basically a function of travel distance, and mean velocity.

4.1.1
Molecular Diffusion

Thermal energy causes an irregular motion of fluid molecules which depends on the mean distance ℓ between molecules, on the collision cross-section σ, and on the distribution of the velocity v. The distance ℓ is proportional to $\rho^{1/3}$, σ is proportional to the cross-sectional area of the molecule, and v is restricted by the principle of the equipartition of energy which demands $\langle v^2 \rangle = 3kT/m$, with velocity v, Boltzmann constant k, temperature T, and mass m of the molecule, or indeed of any particle of interest [*Landau and Lifschitz* 1984, § 29]. Clearly, thermal motion will also affect dissolved substances and small floating particles. Collisions with the fluid molecules will force them into very irregular trajectories (Figure 4.1). These were first reported by *Brown* [1828] who studied pollen grains in water. However, an explanation for the erratic movement in the still water had to await the seminal work of *Einstein* [1905] which subsequently became a cornerstone for the atomic concept of matter [*Renn* 2005].

To focus on the essentials, we consider a gas where the molecules move freely between collisions and we study the movement of some floating particle. The macroscopic motion of the gas shall be negligible. We only consider one component of the motion, along the x-axis. At $t = 0$, the particle is located at $x = 0$. After collision i, it moves by distance Δx_i during time $\Delta\tau_i$ before it collides again. We call these the space and the time increment, respectively, and we notice that both are random variables. Provided the gas is at rest and temperature, pressure, and composition are constant, the statistical properties of the increments will also be constant both in space and time. Furthermore, the expectation value of Δx vanishes, i.e., $\langle\Delta x\rangle = 0$. Finally, we may assume successive increments to be statistically independent due to the irregularity of thermal motion and to the large number of gas particles. This is the prerequisite for the CLT, the central limit theorem (A.24). It ascertains that the position of the particle after a large number n of collisions becomes a Gaussian pdf with expectation 0, since $\langle\Delta x\rangle = 0$, and with the variance equal to n times the variance of the spatial increment. Hence, time and variance of the position become

$$t_n = \sum_{i=1}^{n} \Delta\tau_i =: n\overline{\Delta\tau} \quad \text{and} \quad \sigma_n^2 = \sum_{i=1}^{n} \langle\Delta x_i^2\rangle =: n\overline{\Delta x^2} \ , \tag{4.1}$$

respectively, where the bar indicates the average in time. For sufficiently large values of n, the time averages approaches the averages over a correspondingly large ensemble. Hence, $\overline{\Delta\tau} \to \langle\Delta\tau\rangle$, the mean time between collisions, and $\overline{\Delta x^2} \to \langle\Delta x^2\rangle$, the variance of the free path. This leads to

$$\sigma_n^2 = n\langle\Delta x^2\rangle = t_n \frac{\langle\Delta x^2\rangle}{\langle\Delta\tau\rangle} =: 2D_m t_n \ , \tag{4.2}$$

where the diffusion coefficient D_m is defined following *Einstein* [1905]. Notice that D_m is a macroscopic parameter that depends on properties of the gas and of the diffusing particle but not on time.

With (4.2), we finally obtain for the location of the diffusing particle that passed through $(x, t) = (0, 0)$ the pdf

$$p(x;t) = \frac{1}{\sqrt{4\pi D_m t}} \exp\Bigl(-\frac{x^2}{4D_m t}\Bigr) \tag{4.3}$$

and, in d dimensions, where the components of the motion are independent because of the principle of the equipartition,

$$p(\mathbf{x};t) = [4\pi D_m t]^{-d/2} \exp\Bigl(-\frac{\mathbf{x}\cdot\mathbf{x}}{4D_m t}\Bigr) \ . \tag{4.4}$$

Up to now, we studied the most simple case where we could work with well-defined increments between collisions. The situation is more complicated for diffusion in liquids where there is a continuous interaction due to the much higher density. However, the crucial tool is the CLT and thus the existence of a time increment $\Delta\tau$ such that the respective space increments Δx are statistically independent. Again due to the irregularity of thermal motion, this prerequisite can be ascertained also for liquids. Ideally, the diffusion coefficient D_m is calculated according to (4.2), even for complicated interactions. Alternatively, it may be determined experimentally.

We notice that the approach outlined here can be applied to all phenomena that may be decomposed into a sum of statistically independent increments, even to non-physical processes like for instance to stock markets. As a nod to the original observer this entire class of phenomena is referred to as *Brownian motion*.

Brownian Motion and Molecular Diffusion Molecular diffusion was described well before its fundamental nature was established. In analogy to earlier work by *Fourier* on heat conduction, *Fick* [1855] postulated that the negative gradient of the concentration C was the driving force of the diffusive mass flux $\mathbf{j}_s$ (mass per unit area and unit time). He further proposed a linear relation between flux and driving force,

$$\mathbf{j}_s = -D_m \nabla C \ , \tag{4.5}$$

which is now referred to as *Fick's flux law* or as Fick's first law. The constant of proportionality, D_m, is the coefficient of molecular diffusion. Combining this with the conservation of mass,

$$\partial_t C + \nabla \cdot \mathbf{j}_s = 0 \ , \tag{4.6}$$

he arrived at the *diffusion equation*

$$\partial_t C - D_m \nabla^2 C = 0 \tag{4.7}$$

which is sometimes referred to as Fick's second law. Direct inspection reveals that (4.4) is the solution of (4.7) for an unbounded space with initial concentration $C(\mathbf{x}; 0) = \delta(\mathbf{x})$.

Apparently, there exists a close relation between concentration distributions $C^\delta(\mathbf{x}; t)$ that result fom an initially sharp pulse

$$C^\delta(\mathbf{x}; 0) = m_0 \delta(\mathbf{x}) \ , \tag{4.8}$$

where m_0 is the total solute mass, and the pdf $p(\mathbf{x}; t|0; 0)$ for the transition $0 \to \mathbf{x}$ of a particle between times 0 and t. Indeed, expressing concentration and mass by numbers of molecules, the concentration distribution after time $t > 0$ is found as $C^\delta(\mathbf{x}; t) = m_0 p(\mathbf{x}; t|0; 0)$, hence

$$p(\mathbf{x}; t|0; 0) = \frac{C^\delta(\mathbf{x}; t)}{m_0} \ . \tag{4.9}$$

For a uniform fluid and a stationary flow field, the transition probability only depends on differences in space and time but no more on absolute locations. Hence, $p(\mathbf{x}_1; t_1|\mathbf{x}_0; t_0) = p(\mathbf{x}_1 - \mathbf{x}_0; t_1 - t_0|0; 0)$ which may be abbreviated as $p(\mathbf{x}; t)$. We call this the *travel distance pdf*.

4.1.2
Taylor-Aris Dispersion

Consider a cylindrical tube with stationary flow. The velocity field is given by (2.33), the law of Hagen-Poiseuille, which we write as

$$v_x(r) = 2\overline{v}\left[1 - \left[\frac{r}{r_0}\right]^2\right] \tag{4.10}$$

by using the average velocity (2.35). Transport of solutes in this flow is by two mechanisms, convection with the flow field and molecular diffusion within the fluid. During the short time interval Δt, the average movement of a particle is

$$\Delta x^{\text{conv}} = \overline{v}\Delta t \tag{4.11}$$

by convection and, using (4.2),

$$\Delta x^{\text{diff}} = \sqrt{2D_m \Delta t} \tag{4.12}$$

by diffusion. Obviously, the detailed movement of a single particle is quite complicated as diffusion leads to transitions between flow lines together with the associated changes in convection velocity. To simplify matters, we introduce two time scales. The *longitudinal time scale*

$$\tau_\ell := \frac{2D_m}{\overline{v}^2} \tag{4.13}$$

is characteristic for the transition from diffusion- to convection-dominated longitudinal transport in that $\Delta x^{\text{conv}}(\tau_\ell) = \Delta x^{\text{diff}}(\tau_\ell)$. The *transverse time scale*

$$\tau_t := \frac{2r_0^2}{D_m} \tag{4.14}$$

gives the characteristic time for a solute particle to traverse the cross-section by molecular diffusion. Using these two times, we define the dimensionless quantity

$$\text{pe} := \sqrt{\frac{\tau_t}{\tau_\ell}} = \frac{r_0\overline{v}}{D_m} \tag{4.15}$$

which is the *microscopic Peclet number*. Transport is dominated by convection if $\text{pe} \gg 1$, hence $\tau_t \gg \tau_\ell$. This means that there exists a time interval for which the trajectory of a solute particle may be described by a slightly perturbed streamline. Notice that all radial features of the initial concentration distribution disappear on a time scale of τ_t.

We consider an initial concentration distribution which is radially uniform and highly localized axially, i.e., $C(r, x; 0) = m_0\delta(x)/[\pi r_0^2]$, and, for convection-dominated transport, distinguish three transport regimes (Figures 4.2–4.3):

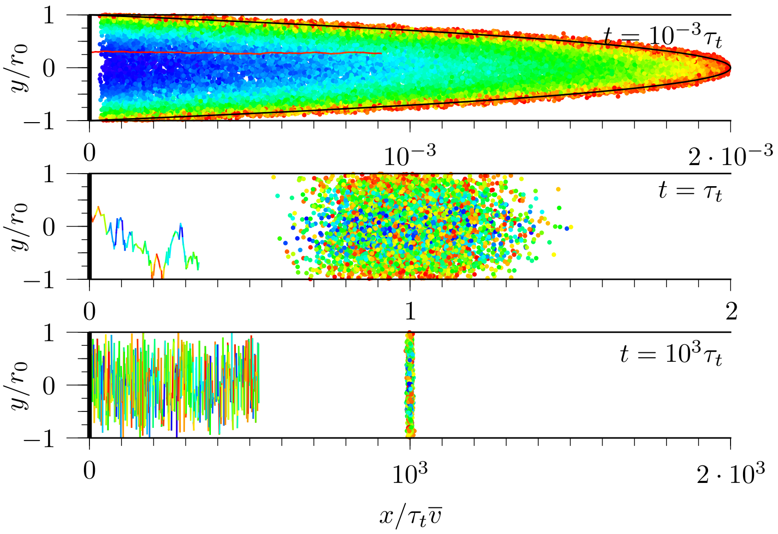


Figure 4.2. Taylor-Aris dispersion with $\mathrm{pe} = 10^6$. At time $t = 0$, all particles are uniformly distributed over the cross-section at $x = 0$. Their positions along the tube are shown for three different times: $t = 10^{-3}\tau_t$, where the influence of Brownian motion is small **(top)**, $t = \tau_t$ for the complicated transitional regime **(middle)**, and $t = 10^3\tau_t$, where the particles have traversed the laminar flow field many times **(bottom)**. The black parabola in the top frame marks the intersection of the cutting plane with the paraboloid (4.16), where all particles would be located if Brownian motion was negligible. The irregular lines show the first half of some particular trajectory. The color of a particle indicates its distance from the cutting plane along the axis. The probability distribution of the particles along the axis is shown in Figure 4.3.

Short-Time Limit For $\tau_\ell \ll t \ll \tau_t$ diffusion may be neglected and the concentration can be approximated by the thin surface given by

$$x(r;t) = v(r)t = 2\overline{v}t\left[1 - \left[\frac{r}{r_0}\right]^2\right] . \tag{4.16}$$

Intermediate Times For $t \approx \tau_t$ the situation becomes rather complicated since diffusion leads to strong but still incomplete mixing.

Long-Time Limit For $t \gg \tau_t$, mixing is complete and we may estimate the concentration distribution with the CLT in analogy to the treatment of Brownian motion. To this end, consider the increment Δx for time intervals τ. Denote the mean velocity of a solute particle along its trajectory by

$$v_\tau := \frac{1}{\tau}\int_0^\tau v\big(r(\tau')\big)\,\mathrm{d}\tau' , \tag{4.17}$$

where $r(\tau)$ is the radial distance from the cylinder's axis. Notice that this is a complicated function: It describes Brownian motion in a circular area, and

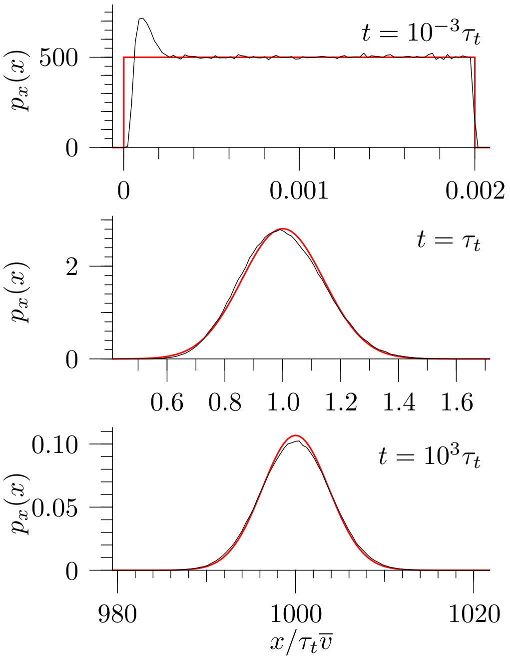


Figure 4.3. Probability distribution function $p_x(x)$ for particles along the axis at the times shown in Figure 4.2. Without molecular diffusion, the distribution in the top frame would be uniform in the interval $[0, 2\overline{v}t]$ (red curve). Diffusion is most manifest for particles with extreme velocities, e.g., at the low end, where diffusion brings particles from near the wall ($v \approx 0$) to regions with higher velocities and thereby leads to the observed peak. At intermediate times $t \approx \tau_t$, the distribution is already nearly Gaussian with only a little bit of asymmetry. The red curves for the lower two graphs are Gaussians that have been fitted to the data. Notice how the relative spread of the distribution decreases with increasing time.

we only know it statistically. The same is then true for $v(r(\tau))$, of course. Hence, v_τ and $\Delta x = \tau v_\tau$ must be considered as random variables. Next, we calculate the expectation of Δx for the ensemble of particles as

$$\langle \Delta x \rangle = \tau \langle v_\tau \rangle = \tau \overline{v} \ , \tag{4.18}$$

where the last equality reflects the fact that diffusion samples the cross-sectional area uniformly. If the initial distribution of particles is not uniform, this is only correct after an initial relaxation time of order τ_t during which the deviations diffuse away. Finally, we estimate the form of the variance of Δx as

$$\mathrm{var}(\Delta x) = \tau^2 \mathrm{var}(v_\tau) \propto \tau^2 \overline{v}^2 \ , \tag{4.19}$$

where the proportionality arises from $v(r) \propto \overline{v}$. The constant of proportionality depends on τ but approaches a constant value as τ increases. We choose $\tau = \tau_t$, define the effective dispersion coefficient

$$D_{\mathrm{eff}} := \frac{\mathrm{var}(\Delta x)}{2\tau_t} \propto \frac{\overline{v}^2 r_0^2}{D_m} = D_m \, \mathrm{pe}^2 \tag{4.20}$$

in analogy to (4.2), and find, by the CLT, for $t \gg \tau_t$ the travel distance pdf

$$p(x;t) = \frac{1}{\sqrt{4\pi D_{\text{eff}} t}} \exp\Bigl(-\frac{[x-\overline{v}t]^2}{4D_{\text{eff}} t}\Bigr) \ . \tag{4.21}$$

This represents a Gaussian pulse which moves with constant velocity $\overline{v}$ (convection) and whose variance increases proportionally to t (dispersion). Comparing (4.21) with (4.3), we recognize that dispersion is formally identical to molecular diffusion. This is expected since both result from the CLT. Notice however, that dispersion results from the spatial variability of the flow field whereas diffusion originates in thermal motion.

The process discussed so far – transport in a cylindrical capillary with laminar flow in the limit of high Peclet numbers – is referred to as *Taylor dispersion*. In his original work, *Taylor* [1953] also gave the value for the missing constant of proportionality and wrote (4.20) as $D_{\text{eff}} = \overline{v}^2 r_0^2/[48D_m]$. *Aris* [1956] extended this to the general case of arbitrary Peclet number and arbitrary cross-section, the so-called *Taylor-Aris dispersion*, and obtained

$$D_{\text{eff}} = D_m + \alpha\frac{\overline{v}^2 r_0^2}{D_m} = D_m\bigl[1 + \alpha\,\text{pe}^2\bigr] \ , \tag{4.22}$$

where α is a dimensionless constant that depends on shape and choice for the characteristic extent r_0 of the cross-sectional area.

The formal analogy between Taylor-Aris dispersion and molecular diffusion can also be extended to the flux law. We formulate it as the sum of the convective flux $\overline{v}C$ that results from the transport of solutes with the mean flow and of the dispersive flux $-D_{\text{eff}}\partial_x C$, hence

$$j_s = \overline{v}C - D_{\text{eff}}\partial_x C \ . \tag{4.23}$$

Inserting this into the one-dimensional form of the mass balance (4.6) then yields

$$\partial_t C + \overline{v}\partial_x C - D_{\text{eff}}\partial_{xx} C = 0 \ , \tag{4.24}$$

a one-dimensional instance of the *convection-dispersion equation*. One finds by inspection that (4.21) is a solution of (4.24) for the case of an unbounded medium with initial condition $C(x;0) = \delta(x)$, and the discussion of the relation between concentration and probability density applies again.

Apparent and Effective Transport Parameters The two limiting regimes of short and long times – the near- and the far-field as we will call them later – lead to characteristically different evolutions for the statistical moments of the travel distance pdf.

In the *far-field*, the travel distance pdf is given by the Gaussian function (4.21) from which we directly read the first two moments

$$\langle x(t)\rangle = v_{\text{eff}} t \qquad \text{and} \qquad \text{var}(x(t)) = 2D_{\text{eff}} t \ , \tag{4.25}$$

where $\overline{v}$ is replaced by v_{eff} in favor of a uniform notation. Notice that the moments can be easily calculated, e.g., from an appropriate tracer experiment, and may thus be used to obtain the effective transport parameters as

$$v_{\text{eff}} = \mathrm{d}_t\langle x(t)\rangle \qquad \text{and} \qquad D_{\text{eff}} = \frac{1}{2}\mathrm{d}_t\mathrm{var}(x(t)) \ . \tag{4.26}$$

This is the so-called *method of moments*. We notice in passing that the definitions

$$v_{\text{eff}}^{\text{glob}} = \frac{\langle x(t)\rangle}{t} \qquad \text{and} \qquad D_{\text{eff}}^{\text{glob}} = \frac{\mathrm{var}(x(t))}{2t} \tag{4.27}$$

would work as well and, in the far-field of Taylor-Aris dispersion, would obviously produce the same numbers as (4.26) since both $\langle x(t)\rangle$ and $\mathrm{var}(x(t))$ are linear homogeneous functions. If this is no more the case, i.e., if the underlying process is no more an effective convection-dispersion, we may still define parameters v and D according to either (4.26) or (4.27). However, we will then call them *apparent* local and global, respectively, since they are merely an alternative description of the first two statistical moments of the travel distance pdf. They must not be interpreted as objective transport parameters. This is discussed further in the following.

The *near-field* of Taylor-Aris dispersion is characterized by the constant velocity along each particle's trajectory. Let $g_v(v)$ be the pdf of the velocity in an orthogonal cross-section of the capillary. Recall that the capillary is cylindrical, but not necessarily circular, and that the flow is laminar. Hence $g_v(v)$ and v are both constant along the axis and $x(t) = vt$. Thus

$$\langle x(t)\rangle = \langle v\rangle t \qquad \text{and} \qquad \mathrm{var}(x(t)) = \mathrm{var}(v)t^2 = \alpha\langle v\rangle^2 t^2 \tag{4.28}$$

where $\alpha := \mathrm{var}(v/\langle v\rangle)$ is some constant that depends only on the geometry of the flow domain. The moments of v then become

$$\langle v\rangle = \int_\Omega v g_v(v)\,\mathrm{d}v \qquad \text{and} \qquad \mathrm{var}(v) = \int_\Omega [v - \langle v\rangle]^2 g_v(v)\,\mathrm{d}v \ , \tag{4.29}$$

where Ω is the capillary's cross-section. Comparing (4.28) with (4.25), we let $v_{\text{eff}} = \langle v\rangle$ and recognize that the two regimes, the near- and the far-field, are indistinguishable when only the first moments are considered. However, the second moments evolve differently, proportionally to t in the far-field and proportionally to t^2 in the near-field. We will encounter this general behavior over and over again, also at larger scales, and employ it for the operational definition

near-field	stochastic convection process (SC)	$\mathrm{var}(x(t)) \propto t^2$
far-field	convection-dispersion process (CD)	$\mathrm{var}(x(t)) \propto t$

For the case of a circular cross-section, the velocity is given by Hagen-Poiseuille's law as $v(r) = 2\overline{v}[1 - [r/r_0]^2]$ and (4.28) yields

$$\langle x(t)\rangle = \overline{v}t \qquad \text{and} \qquad \mathrm{var}(x(t)) = \frac{1}{3}\overline{v}^2 t^2 \ . \tag{4.30}$$

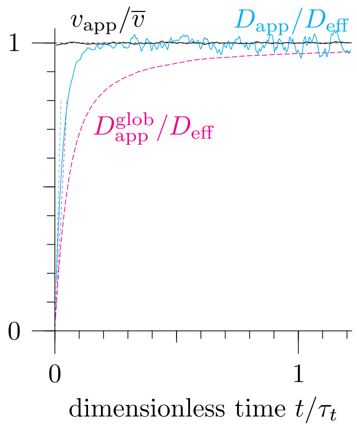


Figure 4.4.
Temporal evolution of apparent parameters v_app (black) and D_app (cyan) estimated from distribution of 10^6 particles in Taylor-Aris dispersion as shown in Figure 4.2. For comparison, the global apparent dispersion coefficient $D^\mathrm{glob}_\mathrm{app}$ (thick dashed magenta) and the approximations for D_app in the limit of short-times (thin dashed) are shown.

Calculating the apparent dispersion coefficient with (4.26) and (4.27) yields

$$D_\mathrm{app} = \frac{1}{3}\overline{v}^2 t \qquad \text{and} \qquad D^\mathrm{glob}_\mathrm{app} = \frac{1}{6}\overline{v}^2 t\ , \tag{4.31}$$

respectively. The two values are by a factor of two different. More importantly – and in contrast to the far-field – they are not constant in time, hence are not inherent transport parameters of the capillary.

The evolution of the apparent parameters for the transition from the near- to the far-field based on the numerical simulation of Taylor-Aris dispersion is shown in Figure 4.4. Apparently, the estimate of D_app is still rather noisy, despite the large number of particles and the good temporal resolution. In practical applications, for instance in a tracer experiment at the field scale, concentration or flux measurements often come with large errors which makes estimates of D_app rather noisy and may even lead to negative values at some points. Such data motivate the use of the more robust $D^\mathrm{glob}_\mathrm{app}$ instead of D_app even though Figure 4.4 illustrates that we may expect significant differences between the two.

4.1.3
Dispersion in Pore-Space

Consider a network of flow channels that are interconnected at discrete junctions. Within a flow channel, transport is a Taylor-Aris process that is generalized such that it includes cross-sections that vary along the axis. On top of the dispersion within each channel, a qualitatively new mixing process then comes into play with the junctions. Depending on the details of the flow within a junction, mixing ranges from purely diffusive to dominantly convective (Figure 4.5). For a uniform network of flow channels, we again expect a convection-dispersion process for the far-field. Interestingly, we will find that the two extremes of local mixing in the junctions lead to

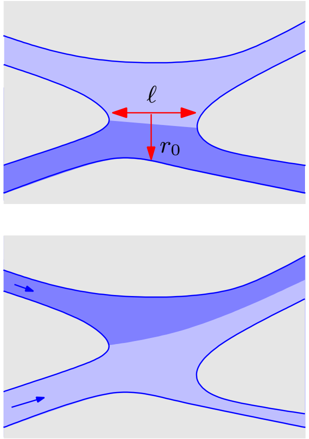


Figure 4.5.
Mixing in junction of pore network ranges from purely diffusive (upper) to dominantly convective (lower), depending on details of the flow. The efficiency of diffusive mixing in the junction depends on the contact time, hence on the velocity. In contrast, convective mixing is constant as long as the flow stays laminar.

characteristically different dependencies of D_{eff} on the microscopic Peclet number pe, hence on the mean velocity. To this end, we will again follow a qualitative line analogous to the one used for understanding Taylor-Aris dispersion. Again, it will provide us with the correct relations up to some constant factor.

Diffusive Mixing Let ℓ be the characteristic length of the junction in the direction of the mean flow and let r_0 be the characteristic radius perpendicular to it. The time available for diffusive mixing is then of order $\ell/\overline{v}$, where $\overline{v}$ is the mean velocity in the junction. The fraction κ of the flow that gets mixed is then of order

$$\kappa = \frac{2D_m t}{\pi r_0^2} = \frac{2D_m \ell}{\pi r_0^2 \overline{v}} = \alpha \frac{\ell}{r_0} \frac{1}{\text{pe}} \ , \tag{4.32}$$

where $\text{pe} = r_0\overline{v}/D_m$ is the microscopic Peclet number introduced in (4.15) and constant $\alpha = 2/\pi$ for a circular cross-section and some other value for another shape.

Next consider the characteristic distance Λ between junctions and notice that $1/\kappa$ is number of junctions that have to be transversed for complete mixing. Hence the distance $\Lambda r_0 \text{pe}/[\alpha\ell]$ has to be covered which leads to the characteristic mixing time

$$t_{\text{mix}} = \frac{\Lambda}{\alpha} \frac{r_0}{\ell} \frac{\text{pe}}{\overline{v}} = \frac{\Lambda}{\alpha} \frac{r_0^2}{\ell D_m} = \text{const} \ . \tag{4.33}$$

At this point, we invoke (4.28) which states that for a stochastic convection process $\text{var}(x(\tau)) \propto \overline{v}^2\tau^2$. Setting $\tau = t_{\text{mix}}$ and using the CLT then yields for $t \gg t_{\text{mix}}$

$$\text{var}(x(t)) = \frac{t}{t_{\text{mix}}} \text{var}(x(t_{\text{mix}})) = \alpha' t_{\text{mix}} \overline{v}^2 t \tag{4.34}$$

where α' is a new constant and, with (4.26), further to

$$D_{\text{eff}} = \alpha'' \frac{\Lambda}{\ell} \frac{r_0^2}{D_m} \overline{v}^2 \qquad \text{or} \qquad \frac{D_{\text{eff}}}{D_m} = \alpha'' \frac{\Lambda}{\ell} \text{pe}^2 \ , \tag{4.35}$$

where α'' is yet another constant. Comparing this with (4.22), we arrive at the important insight that purely diffusive mixing in an interconnected network of flow channels leads to essentially the same transport regime as Taylor-Aris dispersion in a single capillary.

Convective Mixing Consider a junction where a fraction κ of the entering flow is redirected from one flow channel to another as sketched in Figure 4.5. As long as the flow is laminar the shape of the stream lines is independent of the mean velocity and κ is a constant that only depends on the geometry of the flow domain. In the following we neglect the effect of molecular diffusion on mixing in the junction. Proceeding along the same line as before, we obtain Λ/κ for the distance after which the flow in a particular channel is completely mixed with that of its environment. Here, Λ is again the distance between junctions. The corresponding time then is

$$t_{\text{mix}} = \frac{\Lambda}{\kappa \overline{v}} = \frac{\Lambda r_0}{\kappa D_m \text{pe}} \ . \tag{4.36}$$

Notice that now, in contrast to the purely diffusive mixing, the mixing length is constant while the mixing time depends inversely on the mean velocity. With (4.34) and (4.26) we finally obtain

$$D_{\text{eff}} = \beta \Lambda \overline{v} \qquad \text{or} \qquad \frac{D_{\text{eff}}}{D_m} = \beta \frac{\Lambda}{r_0} \text{pe} \ , \tag{4.37}$$

where β is a constant that also includes κ. Notice that the second expression is only useful to put D_{eff} into perspective to the molecular process: diffusion has no role in this mixing process. Correspondingly, the dependence on r_0 is artificial and merely gives the reference for the Peclet number.

Apparently, the linear dependence of D_{eff} on $\overline{v}$ indicates a dispersion process that is fundamentally different from Taylor-Aris. It thus gets a new name and is call *hydromechanic dispersion*. The constant of proportionality between the two quantities is the *dispersivity* λ, hence

$$D_{\text{eff}} = \lambda \overline{v} \ . \tag{4.38}$$

It has the dimension of a length and gives the characteristic extent of the relevant mixing structures.

4.2
Transport in Porous Media

We migrate from a description at the pore space to a continuum description at the scale of the porous medium as a whole by starting out along the lines of

Section 3.3. As a first step, we will invoke appropriate REVs implicitly, define concentrations as the relevant state variables, and formulate the conservation of mass, thereby introducing the macroscopic solute flux. Before the next step – coming up with an empirical flux law – we recall the discussion of Taylor-Aris dispersion in Section 4.1.2. It revealed that only the asymptotic regimes in the near- and in the far-field lead to simple formulations. Intermediate regimes strongly depend on details of the transport process. This contrasts with fluid flow, where we have the luxury of the unique flux law (3.33). The fundamental reason for this difference is that it is immaterial where some fluid element is coming from, all consist of the same fluid. This is quite different for solute transport. Depending on their origins, fluid elements carry different solute masses. The mixing with adjacent fluid elements – molecular diffusion in Taylor-Aris dispersion, mixing at pore junctions in porous geometries – determines the characteristic time or distance over which the initial mass remains in the fluid element, hence the "memory" of the process, or the transition from the near- to the far-field. We notice that this general situation occurs in all transport processes, for instance also in heat transport. In view of these difficulties, relations between concentration and solute flux that are independent of the specific regime will be considered before formulating the empirical flux law and the dynamics of solute transport. Finally, we look into popular parameterizations of the ensuing material properties, in particular the effective dispersion tensor D_{eff}, and into some experimental evidence.

4.2.1
State Variables

The state of a porous medium with respect to solute transport is described completely by the spatial distribution of the solute concentration. Let C_w^μ be the microscopic concentration, the concentration in the water phase at the pore-scale. Consider a volume V of the porous medium and denote the region occupied by the water phase with V_w (Figure 3.15 on page 47). The solute mass in V then is

$$\int_{V_w} C_w^\mu \, \mathrm{d}V = \|V_w\| \langle C_w^\mu \rangle_w \ , \tag{4.39}$$

which evidently is a macroscopic quantity. Referring this mass to a macroscopic volume produces the desired concentration. Since there are two useful volumes, $\|V\|$ and $\|V_w\|$, we obtain two concentrations: the *total concentration*

$$C_t := \frac{\|V_w\|}{\|V\|} \langle C_w^\mu \rangle_w = \theta \langle C_w^\mu \rangle_w \ , \tag{4.40}$$

where θ is the volumetric water content, and the *concentration in the water phase*,

$$C_w := \langle C_w^\mu \rangle_w \ . \tag{4.41}$$

We notice that the quantities C_w and C_t are well-defined if, and only if, volume V is an REV simultaneously for the pore-space, for the water content, and for the solute concentration. This in particular implies that the microscopic quantities θ^μ and C_w^μ are in equilibrium with respect to the macroscopic state given by θ and C_w.

Apparently, C_w and C_t are related to each other by

$$C_t := \theta C_w \ . \tag{4.42}$$

This is only true, however, if the solute is present exclusively in the water phase. In the general case, solute molecules may adsorb at the soil matrix or they may transfer between possible other phases, most importantly a gaseous phase or another immiscible liquid. Then, the total concentration becomes

$$C_t = \sum_i \theta_i C_i \ , \tag{4.43}$$

where θ_i is the volume fraction of phase i and C_i the solute concentration per unit volume of phase i. For a complete description, a set of equations has to be supplied together with (4.43) in order to describe the exchange between the various phases.

We mention that C_w and C_t as defined in (4.40) and (4.41) are sometimes referred to as *resident concentrations* and contrasted to the so-called flux concentration $C^f := j_s/j_w$, where j_s and j_w are the mass flux of solute and the volume flux of water, respectively [e.g., *Kreft and Zuber* 1978]. Resident concentrations are obtained when a porous medium is sampled, whereas the flux concentration would be measured in the outflow from the porous medium. We will not entertain the notion of a flux concentration, however, and always operate directly with the corresponding fluxes.

4.2.2
Mass Balance

The conservation of solute mass is formulated in complete analogy to the conservation of water mass in Section 3.3.4. Let $\mathbf{j}_s^\mu = \mathbf{v}^\mu C_w^\mu - D_m^\mu \nabla C_w^\mu$ be the solute mass flux at the pore-scale. The first term on the right accounts for convection with the local velocity $\mathbf{v}^\mu(\mathbf{x})$, a function which is highly variable in space. The second term represents molecular diffusion, for which we already notice that the coefficient D_m^μ will be smaller near a boundary than within the fluid. With this, the conservation of solute mass for the macroscopic volume V with water phase V_w becomes

$$\partial_t \int_{V_w} C_w^\mu \,\mathrm{d}V = -\int_{\partial V_w} \mathbf{j}_s^\mu \cdot \mathrm{d}\mathbf{A} = -\int_{V_w} \nabla \cdot [\mathbf{v}^\mu C_w^\mu - D_m^\mu \nabla C_w^\mu] \,\mathrm{d}V \ , \tag{4.44}$$

where Gauss' theorem was invoked for the second equality. Choosing V as a fixed volume, the divergence may be pulled out of the integral. Further choosing V as an REV with respect to water content θ, hence $\|V_w\|/\|V\| = \theta$, and invoking the mean value theorem leads to

$$\partial_t[\theta C_w] + \nabla \cdot [\theta \langle \mathbf{v}^\mu C_w^\mu \rangle_w] - \nabla \cdot [\theta \langle D_m^\mu \nabla C_w^\mu \rangle_w] = 0 \ , \tag{4.45}$$

where $C_w := \langle C_w^\mu \rangle_w$ is the average concentration in the water phase of volume V.

We first look into the second term, whose main aspect is $\langle \mathbf{v}^\mu C_w^\mu \rangle_w$, a complicated correlation between microscopic velocity and concentration. To this end, decompose the velocity $\mathbf{v}^\mu(\mathbf{x}) = \mathbf{v} + \mathbf{v}'(\mathbf{x})$ into constant mean $\mathbf{v} = \langle \mathbf{v}^\mu \rangle_w$ and fluctuation $\mathbf{v}'(\mathbf{x})$ with $\langle \mathbf{v}' \rangle_w = 0$ (Reynolds decomposition), and similarly the concentration, $C_w^\mu = C_w + C_w'$. With this, the second term in (4.45) becomes

$$\nabla \cdot [\theta \langle \mathbf{v}^\mu C_w^\mu \rangle_w] = \nabla \cdot [\ \underbrace{\theta \mathbf{v} C_w}_{\text{convection}} \ + \underbrace{\theta \langle \mathbf{v}' C_w' \rangle_w}_{\substack{\text{hydromechanic} \\ \text{dispersion}}} \ ] \ . \tag{4.46}$$

We easily recognize the term $\theta \mathbf{v} C_w$ as the macroscopic convective flux, i.e., the volume flux $\theta \mathbf{v}$ of water with concentration C_w or, alternatively, the total concentration θC_w that moves with velocity $\mathbf{v}$. The second term, the covariance $\theta \langle \mathbf{v}' C_w' \rangle_w$ between the microscopic fluctuations of velocity and concentration, apparently represents hydromechanic dispersion. This covariance is difficult to express in terms of macroscopic quantities, as we may expect already from the discussion of Taylor-Aris dispersion, where different representations for the near- and for the far-field were introduced, and from studying dispersion in the pore-space, where an undetermined dependence of dispersion on the Peclet number, i.e., on the mean flow velocity, was found. Since progress at this point invariably involves assumptions about the geometry of the pore-space and about the covariance of the microscopic velocity, we will only return to this issue with the heuristic formulation of material properties based on experimental evidence.

Next, we consider the third term of (4.45) and easily associate it with molecular diffusion. We may transform it into a macroscopically useful form by noticing that (i) the microscopic concentration C_w^μ and the microscopic flux $\mathbf{j}_s^\mu$ are in equilibrium with the corresponding macroscopic quantities, i.e., changing the macroscopic quantities leads, at the time scale of interest, to an instantaneous change of the corresponding microscopic quantities, (ii) the transport problem considered here is linear, and (iii) D_m^μ may depend in a complicated way on location $\mathbf{x}$ but is otherwise constant. With this, we may write $\langle D_m^\mu \nabla C_w^\mu \rangle_w = D_m \langle d(\mathbf{x}) c(\mathbf{x}) \rangle_w \nabla C_w$ where d and c are complicated functions that describe the microscopic variation of the molecular diffusion coefficient and of the concentration gradient, respectively. Howsoever complicated these functions may be, they depend only on the geometry of the water phase at the pore-scale. For the water-saturated situation we consider here, their spatial average is thus a constant that may be absorbed into an effective value $D_{\text{eff}}^{\text{diff}}$ of the molecular diffusion coefficient for the porous medium. Hence, the third term in (4.45) may be written as

$$\nabla \cdot [\theta \langle D_m^\mu \nabla C_w^\mu \rangle_w] = \nabla \cdot [\theta D_{\text{eff}}^{\text{diff}} \nabla C_w] \ . \tag{4.47}$$

Inserting (4.46)–(4.47) into (4.45) finally yields for the macroscopic formulation of the conservation of mass

$$\partial_t [\theta C_w] + \nabla \cdot [\theta \mathbf{v} C_w] + \nabla \cdot [\theta \langle \mathbf{v}' C_w' \rangle_w - \theta D_{\text{eff}}^{\text{diff}} \nabla C_w] = 0 \ . \tag{4.48}$$

4.2.3
Empirical Flux Law

With (4.48) and the general formulation $\partial_t C_t + \nabla \cdot \mathbf{j}_s$ for the conservation of solute mass, the empirical flux law becomes

$$\mathbf{j}_s = \theta \mathbf{v} C_w + \theta \langle \mathbf{v}' C_w' \rangle_w - \theta D_{\text{eff}}^{\text{diff}} \nabla C_w \ . \tag{4.49}$$

The correlation term $\langle \mathbf{v}' C_w' \rangle_w$ makes this relation rather unwieldy. However, using the same argument as above for the diffusion term $\langle D_m^\mu \nabla C_w^\mu \rangle_w$, it may be written as $\langle \mathbf{v}'(\mathbf{x}) c(\mathbf{x}) \rangle_w C_w$, where $c(\mathbf{x})$ is again a complicated function, different from the one used before, that describes the spatial distribution of the concentration at the pore-scale. With this, we may deduce the general relation that the solute mass flux $\mathbf{j}_s$ is a linear function of the spatial distribution of the macroscopic concentration in the water phase, $C_w(\mathbf{x})$. To emphasize the latter aspect: $\mathbf{j}_s$ is *not* a function of the total concentration C_t (see Exercise 4.1). Hence we may write the macroscopic empirical flux law as

$$\mathbf{j}_s = \theta \boldsymbol{\mathcal{V}} C_w \ , \tag{4.50}$$

where $\boldsymbol{\mathcal{V}}$ is a linear operator that acts on the spatial coordinate of the macroscopic concentration in the water phase. For some special cases, the relation between $\mathbf{j}_s$ and C_w can be readily calculated.

A trivial case is for negligible velocity, i.e., for Peclet number Pe $:= v\ell / D_{\text{eff}}^{\text{diff}} \ll 1$, where v and ℓ are characteristic velocity and size of the transport domain, respectively. Then $\mathbf{j}_s = -\theta D_{\text{eff}}^{\text{diff}} \nabla C_w$.

For convection-dominated transport with $v\ell / D_{\text{eff}}^{\text{diff}} \gg 1$, we first consider the near-field, approximate it by parallel stream-tubes with constant velocities within each tube, and neglect molecular diffusion. Solute pulses in different stream-tubes are thus independent and dispersion within a stream-tube is negligible (Figure 4.6). The empirical flux law for the near-field may then be written as

$$\mathbf{j}_s = \theta \mathbf{v}(\mathbf{x}_\perp) C_w \ , \tag{4.51}$$

where $\mathbf{x}_\perp$ is the location perpendicular to the direction of the flow. The manifestations of this regime will be elaborated in Section 7.1.2.

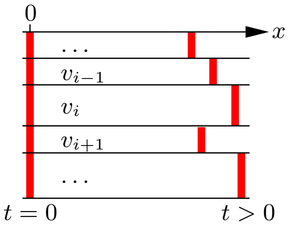


Figure 4.6.
Parallel stream-tubes as an approximation to transport over short distances. The velocity v_i in each tube i is constant and dispersion within a tube is neglected. At time $t = 0$, the solute pulse is located at $x = 0$. It is dispersed according to the distribution of velocities for $t > 0$.

For convection-dominated transport in the far-field of a macroscopically uniform porous medium, we again expect the solute flux to consist of two components – convection and dispersion – with the empirical flux law

$$\mathbf{j}_s = \theta \mathbf{v} C_w - \theta \mathsf{D}_{\text{eff}} \nabla C_w \ , \tag{4.52}$$

where D_{eff} denotes the effective hydrodynamic dispersion tensor. The explicit derivation of the second term, $-\theta \mathsf{D}_{\text{eff}} \nabla C_w$, has been the focus of many studies for porous media, among them *Bear* [1961] and *Scheidegger* [1961], but also for the formally similar transport in turbulent flow [e.g., *Shraiman and Siggia* 2000; *Blackman and Field* 2003].

4.2.4
Dynamics

For the far-field in a macroscopically uniform porous medium (4.48) becomes

$$\partial_t[\theta C_w] + \nabla \cdot [\theta \mathbf{v} C_w] - \nabla \cdot [\theta \mathsf{D}_{\text{eff}} \nabla C_w] = 0 \ , \tag{4.53}$$

which is the *convection-dispersion equation* (CDE) in its general form where θ and $\mathbf{v}$ may be functions of space and time and D_{eff} is in general a function of θ and $\mathbf{v}$.

Some comments are in order, here. (i) In these lecture notes, the subscript in D_{eff} or D_{eff} refers to the fact that in the far-field, dispersion is formally very similar to molecular diffusion even though the underlying mechanisms are quite different. In other literature, one may find the notion $D_{\text{eff}} = \theta D_w$, hence “effective” there refers to the implicit incorporation of the water content θ and D_w refers to the dispersion coefficient. As a mnemonic, “v and D_{eff} are at the same level”, i.e., they both get multiplied by θ when used in a transport equation. (ii) The formulation (4.53) represents the transport of solute that exists exclusively in the water phase, a so-called conservative solute. (iii) At the scale of current interest, at the transition from the pore-scale to the continuum of a macroscopically uniform medium, the convection-dispersion equation is a very good description, even though it only applies at the far-field. The reason for this is that the far-field, at the scale of interest for most soil physical issues, is rather near, say a few tens of grains corresponding to a few millimeters. Molecular diffusion at these scales is rather fast, a typical diffusion coefficient of $D_m = 10^{-10}$ m^{2}s^{-1} leads to typical times of a few hours, which is very short for most, but not all, processes of interest. However, the issues discussed here are generic and not restricted to the transition from the pore-scale and they become dominant in natural porous formation, soil or aquifers, which typically exhibit heterogeneities at many scales.

4.2.5
Material Properties

The first step for solving a transport problem, e.g., for the far-field as described by (4.53), is always to solve the underlying flow problem. This may be as simple as stating a constant flux $\mathbf{j}_w$ for stationary flow in a saturated uniform medium or as complicated as calculating a multiphase flow field, for a specific time, in a heterogeneous medium with transient forcing. The solution of the flow problem yields the water flux $\mathbf{j}_w$ and the volumetric water content θ, both for the flow domain and for the time domain of interest. The material properties with respect to transport relate these macroscopic flow quantities to macroscopic transport quantities, for instance to the effective velocity $\mathbf{v}$ and the effective dispersion tensor D_{eff} in the far-field approximation (4.53).

Predicting effective transport properties from “first principles”, from statistical properties of the pore-space geometry and the corresponding fluid flow, is not yet feasible because these two aspects are not yet understood quantitatively. The main reason for this is the pore-space geometry of natural porous materials, which is much more complicated than what is handled by current theoretical approaches. Coming next to the desire of first-principles insight so far are numerical simulations of flow and transport processes that operate on geometries that have been determined with appropriate micro-tomography instruments.

For the time being, until some deeper understanding emanates, we will follow a different, experiment-based approach. We will do this only after looking into some of the popular parameterizations of effective transport properties.

Effective Velocity Obtaining the effective velocity appears straightforward at first sight,

$$\mathbf{v} = \frac{\mathbf{j}_w}{\theta} \tag{4.54}$$

from (3.31), and we will mostly use exactly this equality. It is important to notice, however, that in general not all of the water phase will contribute uniformly to solute transport, as has been assumed for (4.49). An important instance of this is the interaction between surface charges of the porous matrix and ionic solutes. This may either lead to the exclusion of solute molecules from the immediate neighborhood of the matrix surface, thereby to a lower effective water content, and thus to a higher velocity than predicted by (4.54). The converse effect is observed when solute molecule are attracted to the matrix surface. Another cause, often in association with the action of surface charges, is filtration in very fine textured media where large molecules, ions with large hydration shells, or large particles like viruses cannot enter fine pores. In all these cases, the discrepancy arises from the volume fraction relevant for flow being different from that for transport.

A still more interesting situation arises if different parts of the water phase exhibit vastly different equilibration times with respect to transport. This is for instance the case for a granular medium where the grains are porous themselves or for interstitial water that is typical for unsaturated states. For such situations (4.54) is the correct formulation for sufficiently long times but it may be quite wrong for shorter times. Such systems will be studied in greater detail in Section 7.1.5.

Molecular Diffusion Independent of any flow, molecular diffusion is always present as a dispersion processes. In contrast to diffusion in a pure fluid, spreading in a porous medium is limited, however, which leads to a reduction of the effective diffusion coefficient. A popular parameterization are the *Millington-Quirk* models [*Millington* 1959; *Millington and Quirk* 1961]

$$\frac{D_{\mathrm{eff}}^{\mathrm{diff}}}{D_m} = \frac{\theta^{7/3}}{\phi^2} \qquad \text{and} \qquad \frac{D_{\mathrm{eff}}^{\mathrm{diff}}}{D_m} = \frac{\theta}{\phi^{2/3}} \,, \tag{4.55}$$

where ϕ is the porosity and θ the volumetric water content. Notice that in the original publications, expressions for θD_{eff} as used in (4.49) are given. Hence they differ by the factor θ from (4.55). Further notice that, while the first model is used more often, *Jin and Jury* [1996] report that the second one leads to a better agreement with experimental results. We emphasis that (i) all these studies are aimed at *diffusion of gases in porous media* and that the results are just transposed to the case of solute diffusion in the liquid phase and (ii) there exist a number of further parameterizations all with rather weak experimental support.

Hydrodynamic Dispersion For a natural porous medium, the pure dispersion regimes – molecular diffusion, diffusive mixing, and convective mixing (hydromechanic dispersion) – are hardly ever found and we usually encounter a convolute of all three of them. For any particular medium, the relative importance of the different regimes will obviously depend on the mean velocity since mixing in individual junctions will shift towards diffusive mixing as the velocity decreases and hence the time for interaction increases. In the unsaturated case, we further expect a strong dependence of dispersion on the water content since interstitial water, which is separated from the main flow by narrow films, will lead to a broadening of the velocity distribution.

Before looking into the more difficult issues, we notice that the dispersion coefficient is a tensor since it relates two vector quantities to each other, the concentration gradient ∇C and the dispersive solute flux $\mathbf{j}_s^{\mathrm{disp}}$. It reduces to a scalar only in a one-dimensional setting. Intuitively, one may expect $\mathsf{D}_{\mathrm{eff}}$ to be symmetric with one major axis in the direction of the mean flow, the longitudinal component, and the other two perpendicular, the transverse components. In an isotropic medium, the two transverse components are equal because of symmetry. A number of theoretical studies, most prominently *Bear* [1961] and *Scheidegger* [1961], demonstrated that for the case of pure hydromechanic dispersion in an isotropic medium, the ij-component of D_{eff} may indeed be written as

$$D_{ij} = [\lambda_\ell - \lambda_t]\frac{v_i v_j}{|\mathbf{v}|} + \lambda_t|\mathbf{v}|\delta_{ij} \ , \tag{4.56}$$

where λ_ℓ and λ_t are the longitudinal and transverse dispersivity, respectively, $\mathbf{v}$ is the mean velocity with components v_i and v_j in i- and j-direction, respectively, and δ_{ij} is Kronecker's delta. Notice that (4.38) is retrieved as a special case for the longitudinal dispersion. *Bear* [1972] conjectured that this relation may be extended to arbitrary values of the Peclet number by multiplying the right hand side with some function that depends on pe and on the mean geometry of the flow channels.

Experimental Evidence Literally hundreds of experiments have been performed to investigate $\mathsf{D}_{\text{eff}}(\text{pe})$ and the research is ongoing, particularly with respect to multi-scale media. The common trait of these experiments is to (i) apply a tracer, (ii) monitor its transport, and (iii) deduce effective properties from these measurements. Details of course vary greatly.

Tracers These range from simple salts like $CaCl_2$ or $CaBr_2$ with the hydrated Cl^- and Br^- ion as tracers, through dyes like brilliant blue, to instrument-specific probes like gadolinium(III) compounds for MRI (magnetic resonance imaging), and further to more exotic substances like quantum dots or DNA markers. Such tracers are applied as highly localized impulses, lines, sheets, or steps and they enter the porous medium either with the water flux or are deposited as an initial concentration.

Monitoring In earlier times, monitoring was mostly as breakthrough curves at the lower end of some lab column or at some point instrument [*Pfannkuch* 1963], occasionally also by sampling the medium a few times. Modern approaches aim at high-resolution measurements, simultaneously in space and time, which became feasible with the availability of a range of tomographic instruments, from X-ray to MRI, neutron imaging, or positron emission tomography [*Manz et al.* 1999; *Kandhai et al.* 2002].

An example of such a method is PLIF (planar laser-induced fluorescence) with index-matched solids and fluids [*Stöhr* 2003; *Stöhr et al.* 2003]. In this approach, a porous medium is constructed with transparent grains – glass or plexiglass beads – and the fluid is chosen such that it has exactly the same refractive index as the solid. Such a medium is optically uniform and light passes through it with minimal deflection. Hence, an optical tracer, e.g., some fluorescent substance, can be scanned by a laser sheet that illuminates subsequent slices of the medium. Recording the emitted light with a high-resolution camera yields quantitative three-dimensional tracer distributions

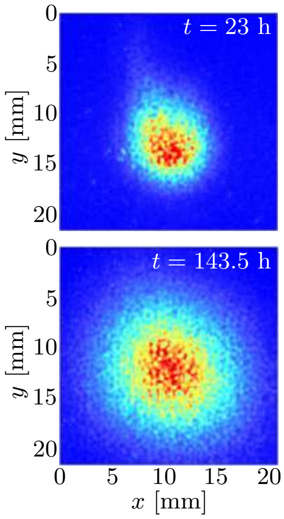


Figure 4.7.
Concentration distribution of a dye tracer (Nile Red) measured with PLIF in a 0.6 mm thick vertical cross-section through a uniform isotropic porous medium. The fluid phase (silicone oil) during this experiment was immobile. At time $t = 0$, the tracer was injected with a syringe and formed a plume of some 5 mm diameter. We expect the concentration C_w in the fluid phase to vary smoothly in space since molecular diffusion rapidly equilibrates local differences. The concentration measured with PLIF corresponds to C_t, however, since it measures the intensity of fluorescent light emanating from a specific volume of porous medium. Since an individual pixel is much smaller than an REV, C_t at this resolution is a spatially highly variable quantity and reflects the grains within the thin visible slice. (Adapted from Figure 8.1 of *Stöhr* [2003])

with a sub-millimeter spatial resolution and a repetition interval of a few seconds. An illustration of the result, for pure molecular diffusion in a porous medium, is shown in Figure 4.7.

Deducing effective properties A still popular approach for estimating transport properties is the method of moments, a particular form of which was introduced with (4.26)–(4.27). It is typically replaced by fitting analytic solutions of the particular transport problem if the data are not complete, e.g., if only part of the breakthrough has been measured. Increasingly, comprehensive numerical inversions are run where details of the experiments can be accounted for that are not accessible to the more traditional methods, i.e., by explicitly accounting for known structures or complicated hydraulic regimes.

Experimental Results While several of the modern experimental variants are capable to yield information on the full dispersion tensor, the majority of the investigations so far focused on the longitudinal component D_ℓ only. Results are often expressed in terms of the microscopic Peclet number $\mathrm{pe} = r_0\overline{v}/D_m$, which was introduce with (4.15).

Experimental as well as numerical studies at the lab scale indicate that $D_{\mathrm{eff}}(\mathrm{pe})$ varies smoothly between the regime of very low values of pe, where essentially molecular diffusion prevails, through the range of intermediate values where D_{eff} is often parameterized as proportional to pe^α with $1 < \alpha < 2$ where both diffusive and convective mixing are relevant, and intimately intermingled, to the final region of very large values, but still laminar flow, where hydromechanic dispersion prevails. Detailed experiments indicate, however, that even for the highest range of Peclet numbers, dispersion does not become purely hydromechanic. There appear to be at least two reasons for this discrepancy: As a first reason, high-resolution experiments show that

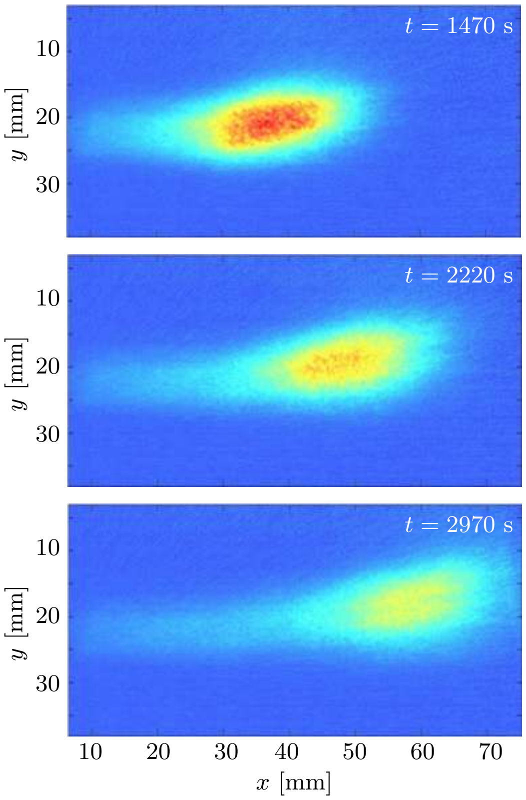


Figure 4.8. Transport of dye tracer (Nile Red) through uniform porous medium with constant flow as seen in a vertical cross-section obtained with PLIF. At time $t = 0$, the tracer was injected as a plume of some 5 mm diameter. Notice that (i) longitudinal dispersion is significantly larger than transversal and (ii) a weak but long tail forms. The former is already predicted by the tensor (4.56), at least qualitatively, while the latter will only be understood with Section 7.1.5. (Adapted from Figure 8.22 of *Stöhr* [2003])

a tracer pulse develops a weak but rather long tail on its way through the porous medium as is illustrated in Figure 4.8. This is explained by tracer particles that diffuse into less mobile regions, which are present in every porous medium, at least as interstitial water next to the grain contacts. As we will see in Section 7.1.5, this tail is expected to disappear eventually, on a time scale that corresponds to the exchange time of the immobile region, and the pulse will approach the expected Gaussian shape. Correspondingly, the true value of D_{eff} will be larger than that estimated from the main pulse alone and it will be reached on a time scale that is large compared to the time scale on which the main pulse, without the tail, evolves into a Gaussian.

A second reason for the deviation from pure hydromechanic mixing is the fact that molecular diffusion is very efficient over short distances. Hence even where convective mixing prevails, diffusion across streamlines will increase the mixing and lead to a stronger dependence on pe. This is illustrated in Figure 4.9 which summarizes experimental results from lab scale experiments. Apparently, they may be described by the heuristic power function

$$D_{\text{eff}}/D_m = \gamma\, \text{pe}^{\alpha} \; . \tag{4.57}$$

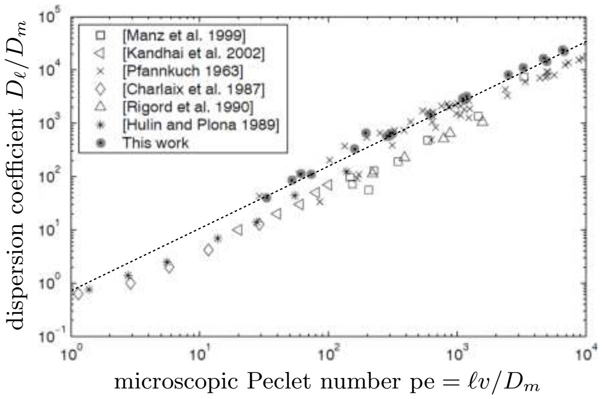


Figure 4.9. Experimentally determined relation between longitudinal dispersion coefficient and Peclet number for unconsolidated porous media. The dashed line is the heuristic function (4.57) with $\gamma = 0.8$ and $\alpha = 1.17$. ("This work" is *Stöhr* [2003]. The figure was adapted from his Figure 8.11.)

There are propositions to describe D_{eff} as a sum of contributions from the individual dispersion processes as $D_{\text{eff}} = \gamma_0 D_m + \gamma_1 \text{pe} + \gamma_2 \text{pe}^2 + \gamma_3 \text{pe} \log(\text{pe})$, where the last term stems from dispersion in the boundary layer adjacent to the solid. We notice that such a model presumes that the individual processes are independent, at least statistically.

Exercises

4.1 Driver for Macroscopic Solute Flux Consider a one-dimensional situation where θ is piecewise constant but jumps at some point and let the water j_w be constant. Consider two configurations, (i) $C_w =$ const and (ii) $C_t = \theta C_w =$ const, calculate the divergence $\partial_x j_s$ for purely convective transport, hence $j_s = \theta v C_w$, and discuss the result. [You may want to extend this to other cases with a known explicit flux law, i.e., pure diffusion, near-field, and far-field.]

5
Groundwater Flow

Groundwater encompasses the part of subsurface water that fills the pore space completely. Formations that carry groundwater are separated into *aquifers*, high permeability hydraulic conductors, and *aquicludes*, low permeability hydraulic isolators. There is no sharp boundary between aquifer and aquiclude and sometimes even an intermediary formation is introduced, the *aquitard*, whose permeability is intermediary between the two.

At a typical site, we find multi-story sequences of aquifers and aquicludes. They originate from contrasting depositional regimes, for instance due to the cycles of ice ages in the past million years, and often have been transformed by various geologic processes. The Upper Rhein valley is an example of such a structure (Figure 5.1). Its tectonic environment is a Graben between Black Forest and Vosges which sank at a rate of 0.2. . . 0.9 mm y^{-1} for the past few million years and continues to do so today. Rhein river continuously filled the deepening Graben with sediment the texture of which varied greatly between glacial and interglacial periods. Coarse textured material like gravel and sand lead to aquifers while very fine textured material, mostly clay, evolved into aquicludes. Part of these layers were eroded again after deposition which led to windows, lenses, and partial layers. Later, differential sinking of the Graben caused the formations to tilt and fracture, with large blocks slipping past each other. All these processes contributed to the rather complicated but typical architecture of this site.

The topmost parts of the subsurface typically belong to the soil water zone where the pore space is filled with temporally and spatially varying fractions of water and air. This is often referred to as *vadose* or *unsaturated zone*. Its thickness varies greatly, from practically 0 in swamps to hundreds of meters in arid regions. This zone is typically followed by the topmost aquifer, the so-called *phreatic* aquifer. It is unconfined, which means that the water table is free to move. The water table is define as the depth where $\psi_m = 0$, hence the pressure in the water phase equals that of the atmosphere. The transition between the vadose zone and the unconfined aquifer is the capillary fringe, i.e., the zone where the matric potential is negative, $\psi_m < 0$, but the pore space is already saturated with water, $\theta = \phi$ (Figure 5.2). The position of the water table may be read from a *piezometer* which is a tube that is open to

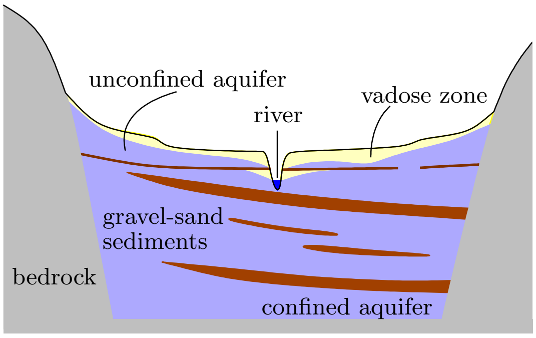


Figure 5.1.
Sketch of a multi-story valley aquifer as it may develop in a tectonic Graben like the Rhein valley. The groundwater stories (blue) are separated by aquicludes (brown) which may be punctured by windows. Height is exaggerated.

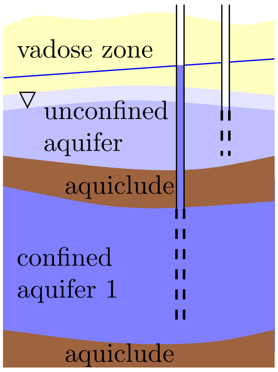


Figure 5.2.
Sketch of two-story groundwater system with piezometers indicating water table of unconfined aquifer and piezometric head (blue line) of confined aquifer. The light blue region between the vadose zone and the unconfined aquifer indicates the capillary fringe.

the atmosphere on one end and to the aquifer that is measured on the other. Notice that the capillary fringe is not captured by this instrument.

A confined aquifer is bounded by two aquicludes. Its thickness thus does not change and it has no free water table. A piezometer shows a water level that is above the lower end of the upper bounding aquiclude. This water level corresponds to the so-called *piezometric head* which we easily associate with the water potential and with the hydraulic head introduced with (3.20). Notice that the piezometric head may be above ground level, in which case one may drill a well from which water will flow freely. Such *artesian wells* are found in many regions for instance in the Great Plains of North America, in the Paris basin, and in the Great Artesian Basin of Australia.

5.1
Dynamics of Flow in Confined Aquifer

We consider a consolidated and confined aquifer. The dynamics of water flow may be formulated by inserting (3.61) and (3.67) into (3.37) which yields the *groundwater equation*

$$S\partial_t p - \nabla \cdot \left[\mathsf{K}[\nabla p - \rho \mathbf{g}]\right] = -\gamma \ , \tag{5.1}$$

where γ is the volume extraction rate, i.e., the volume of water extracted per unit volume of aquifer and unit time. The extraction term accounts for pumping wells, leaking pipes and the like.

We first consider the simple case of a uniform isotropic medium and furthermore neglect gravity. With this, (5.1) simplifies to the diffusion equation

$$\partial_t p - D_p \nabla^2 p = -\frac{\gamma}{S} \,, \qquad D_p := \frac{K}{S} \tag{5.2}$$

for the hydraulic pressure (see Section 4.1.1 for details on diffusion). Here, D_p $[\mathrm{L}^2\mathrm{T}^{-1}]$ is the pressure diffusion coefficient which describes the spreading of an initially localized pressure fluctuation with time.

The fact, that pressure in groundwater obeys a diffusion equation may at first sight astonish. After all, we obtain a wave equation for the pressure in a free fluid, a lake or the atmosphere. However, in porous media, with slow flow described by Stokes' equation (2.27), dissipation is so large that kinetic energy cannot build up. Hence, energy cannot oscillate between kinetic and potential form, which is the prerequisite for a wave.

5.1.1
Stationary Flow

Introducing dimensionless variables through the relaxation time τ of the aquifer and the characteristic length ℓ together with the corresponding differential operators – see (2.21) –, we write (5.1) as

$$\frac{1}{\tau}\partial_{t'} p - \frac{K}{\ell^2 S}\nabla' \cdot \big[\mathsf{K}'[\nabla' p - \ell\rho\mathbf{g}]\big] = -\frac{\gamma}{S} \,, \tag{5.3}$$

where $\mathsf{K} = K\mathsf{K}'$ with $|\mathsf{K}'|$ of order 1. Hence, we define the relaxation time of the aquifer as

$$\tau := \frac{\ell^2 S}{K} = \frac{\ell^2}{D_p} \,. \tag{5.4}$$

Notice that for an external forcing on a time scale $t \gg \tau$, the flow becomes stationary and (5.1) may be simplified to

$$\nabla \cdot \big[\mathsf{K}[\nabla p - \rho\mathbf{g}]\big] = \gamma \,. \tag{5.5}$$

Example: A Typical Sandstone Aquifer In the example on page 73, we found the storage coefficient $S = 1.4 \cdot 10^{-10}$ Pa^{-1}. Assuming an isotropic permeability $k = 10^{-11}$ m^2, which is a rather high value, and a flow domain of interest of 5 km, we find 55 $\mathrm{m}^2\mathrm{s}^{-1}$ for the coefficient of pressure diffusion and 5.3 days for the relaxation time. The approximation of stationary flow – or, equivalently, of incompressible media – is thus permissible if the time scale of external forcing is much larger than a few days. Notice that τ becomes much larger in less permeable or in more compressible formations and that it increases quadratically with the extent of the flow domain.

For isotropic and uniform media, (5.5) reduces to the *Poisson equation*

$$\nabla^2 p = \frac{\gamma}{K} \tag{5.6}$$

and further to the *Laplace equation*

$$\nabla^2 p = 0 \tag{5.7}$$

for a negligible extraction rate.

5.1.2
Large Aquifers

By their nature, aquifers are much more extended horizontally than vertically. While a three-dimensional representation is necessary for local studies, for instance near a pumping well or some contamination source, a two-dimensional representation suffices for most regional studies. An example is the High Plains aquifer in the USA which extends from 97°W to 105°W and from 32°N to 43°N between South Dakota and Texas and covers an area of some $200'000$ km^2. However, its thickness does not exceed 100 m for most of the region although there are patches that are up to 400 m thick. The flow domain is thus essentially two-dimensional and so is the flow field. Also with less extreme ratios between horizontal and vertical dimensions, one often finds that the vertical component of the driving force – the deviation $-\partial_z p + \rho g$ from hydrostatic equilibrium – is negligible compared to the horizontal component $-\nabla_h p$, where $\nabla_h = \{\partial_x, \partial_y\}$ in Cartesian coordinates. Isosurfaces of the water potential are then vertical and (5.1) may be reduced to

$$S_\ell \partial_t p - \mathsf{T} \nabla_h^2 p = -\Gamma \ , \tag{5.8}$$

where S_ℓ, T, and Γ are the vertical integrals over the aquifer's depth ℓ of S, K, and γ, respectively. In the uniform medium considered here and assuming that ℓ is constant, this just corresponds to a multiplication of all the parameters with ℓ. The quantity T is called the *transmissivity*.

We mention that ℓ in general varies in space and thus turns the medium into a heterogeneous one, even if it is uniform in a three-dimensional description.

5.2
Stationary Flow in Uniform Aquifer

A wide selection of solutions for different dimensions, boundary conditions, and initial conditions is available for the Laplace as well as for the Poisson equation. To some extent this is also the case for the more general form

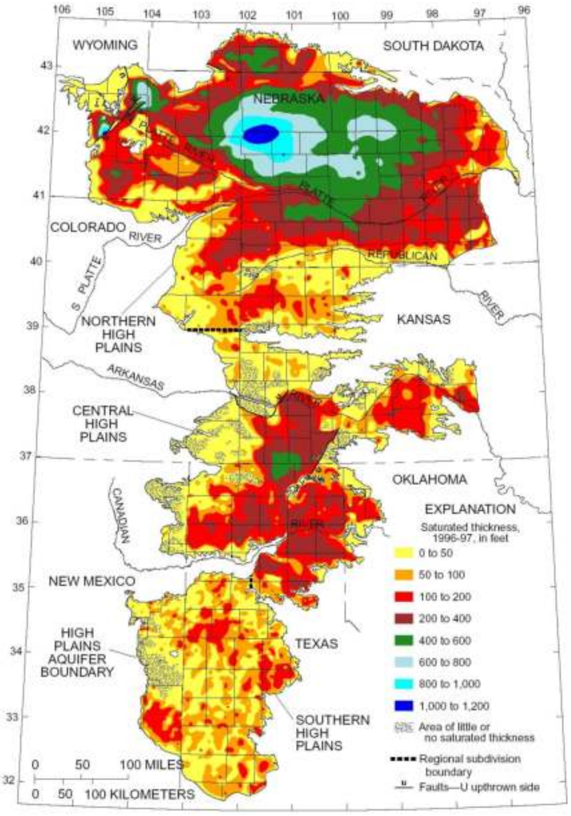


Figure 5.3. Saturated thickness of the High Plains aquifer. [*Figure 4 from McGuire and Fischer [1999]*.]

(5.1). They originate not only from the groundwater literature but also from electrostatic and heat flow problems [e.g., *Bear* 1972; *Carslaw and Jaeger* 1990]. Furthermore, since the equations are linear, they are readily solved directly by using Green's functions, integral transforms, and finally the principle of superposition.

5.2.1
Pumping Well in Regional Flow

We consider a horizontal unbounded aquifer of constant thickness and uniform isotropic properties. From a fully penetrating pumping well located at $\mathbf{x} = 0$ water is extracted at the constant specific rate γ (volume of water per unit time and unit length of well screen). This leads to the two-dimensional problem

$$\mathsf{K}\nabla_h^2 p = \gamma\delta(\mathbf{x}) \ , \qquad \lim_{|\mathbf{x}|\to\infty} \nabla_h p = \mathbf{A} \ , \quad p(r_0) = p_0 \ , \tag{5.9}$$

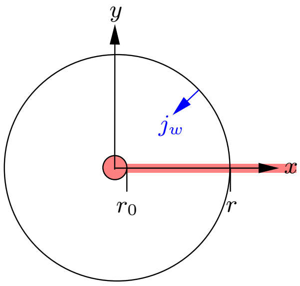


Figure 5.4. Sketch for calculating the pressure field of a single pumping well with radius r_0. The red region is excluded for the later calculation of the stream function (Figure 5.5) such that $\nabla \cdot \mathbf{j}_w = 0$ is satisfied in the remainder.

where $\mathbf{A}$ is the hydraulic gradient far from the pumping well, the so-called regional hydraulic gradient, and p_0 is the pressure at the boundary of the pumping well at distance r_0 from the origin.

This problem is solved most easily by decomposing it into two parts: uniform regional flow that satisfies

$$\mathsf{K}\nabla_h^2 p = 0 \ , \qquad \nabla_h p = \mathbf{A} \ , \quad p(0) = 0 \tag{5.10}$$

and flow towards a pumping well described by

$$\mathsf{K}\nabla_h^2 p = \gamma\delta(\mathbf{x}) \ , \qquad \lim_{|\mathbf{x}|\to\infty} \nabla_h p = 0 \ , \quad p(r_0) = p_0 \ . \tag{5.11}$$

Clearly, (5.9) is solved by the sum of solutions of (5.10) and (5.11), which is an example of the principle of superposition for solutions of linear differential equations.

Uniform Regional Flow The solution of the Laplace equation (5.10) is $p(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$.

Flow towards Single Pumping Well The second problem, (5.11), may be approached by noticing that the flow through each circle around the pumping well equals γ. Together with the radial symmetry we then obtain for arbitrary radius r

$$\gamma = -2\pi r j_w = 2\pi r K \partial_r p \ , \tag{5.12}$$

where Darcy's law was used for the second equality (Figure 5.4). Integration from the well radius r_0 to r then yields

$$p(r) = p_0 + \frac{\gamma}{2\pi K} \log\left(\frac{r}{r_0}\right) \ , \tag{5.13}$$

where p_0 is the pressure at r_0. Notice that this solution has the pressure increasing very slowly, but without bounds as r goes to infinity. This is a consequence of our abstraction of reality: we did not account for recharge through the vertically bounding aquicludes or through some horizontal boundary, e.g., a river, which in reality would eventually supplement the water that is pumped from the aquifer. Still, (5.12) is a reasonable approximation for distances that are not too large.

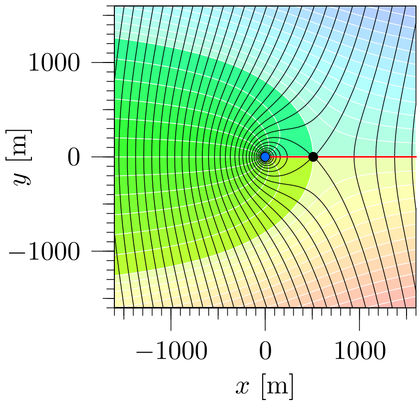


Figure 5.5. Isobars (black lines) and stream function (white lines) for a single pumping well (blue dot) in an otherwise uniform regional flow from the left to the right. The capture region of the well is represented by green hues and the stagnation point, where the velocity vanishes, by the black dot. Notice that the jump in the streamfunction at $y = 0$ for $x > 0$ (red line) equals the flux that is extracted by the pumping well. Parameters are chosen to roughly correspond to the case study in Section 5.6. Specifically: $K = 10^{-3}\ \mathrm{m\,s^{-1}}$, $A = 10^{-2}$, $\gamma = 0.0032\ \mathrm{m^3 s^{-1}}$.

Adding the two solutions finally yields for the solution of (5.9) in Cartesian coordinates

$$p(x,y) = p_0 - Ax + \frac{\gamma}{2\pi K}\log\left(\frac{r}{r_0}\right) , \tag{5.14}$$

where the hydraulic gradient is assumed to be anti-parallel to the x-axis and $r = \sqrt{x^2 + y^2}$. The corresponding flux field may be visualized by the stream function (Section A.1.4 in the appendix). To this end, we calculate the water flux from Darcy's law,

$$\mathbf{j}_w(x,y) = \left\{AK - \frac{\gamma x}{2\pi r^2}, -\frac{\gamma y}{2\pi r^2}\right\} . \tag{5.15}$$

The stream function φ is obtained by integrating $\mathrm{d}\varphi = -j_{w_y}\mathrm{d}x + j_{w_x}\mathrm{d}y$. Recalling that it is only properly defined in regions where $\nabla \cdot \mathbf{j}_w = 0$ we choose a region of interest that excludes the pumping well proper and a small region around the positive x-axis (red area in Figure 5.4). We finally choose $\varphi(-\infty, 0) = 0$ as reference for the stream function to obtain (Figure 5.5)

$$\begin{aligned}\varphi(x,y) &= \int_0^y j_{w_x}(-\infty,\eta)\,\mathrm{d}\eta - \int_{-\infty}^x j_{w_y}(\xi,y)\,\mathrm{d}\xi \\ &= AKy - \frac{\gamma}{2\pi}\arctan\left(\frac{y}{x}\right) + \gamma\left[\mathrm{H}(y) - \frac{1}{2}\right] ,\end{aligned} \tag{5.16}$$

where Heaviside's unit step function $\mathrm{H}(y)$ compensates the discontinuity of $\arctan(y/x)$ at $y = 0$ for negative values of x.

Looking at our solution for the pumping well in a uniform regional flow, we find that the capture zone of the pumping well extends to infinity. Recalling that the flow between two stream lines $\varphi(x,y) = q_1$ and $\varphi(x,y) = q_2$ equals

$q_1 - q_2$ and using that the flow field far away from the well is not disturbed, we obtain for the width ℓ of the capture zone far away from well

$$\ell = \frac{\gamma}{AK} \, . \tag{5.17}$$

Finally, we calculate the stagnation point, the location where the flux vanishes, from solving (5.15) for $j_{w_x} = 0$ and obtain

$$x_s = \frac{\gamma}{2\pi AK} = \frac{\ell}{2\pi} \, . \tag{5.18}$$

We notice in passing that the capture zone is of particular importance for the quality of the groundwater pumped from a well since all contaminants that are resident in this zone or released into it will eventually end up in the extracted water.

5.2.2
Dipole Pumping in Regional Flow

By far the most common pumping wells are single extraction wells, or galleries of them, for water production. However, for the remediation of contaminated groundwater, dipole or eventually multi-pole pumping is more practical. In such an arrangement, water is extracted from the aquifer, treated in order to remove the contaminants, and returned to the aquifer at another location. Pressure and flow fields of such set-ups can be easily composed from the solutions derived in the previous example by using the principle of superposition and (5.13) as a Green's function for a single pumping well. In this way, we write the pressure field that results from an extraction well at $(-x_0, 0)$ and a recharge well at $(x_0, 0)$ as

$$p(x,y) = -Ax + \frac{\gamma}{2\pi K}\left[\log\left(\frac{\sqrt{[x+x_0]^2+y^2}}{r_0}\right) - \log\left(\frac{\sqrt{[x-x_0]^2+y^2}}{r_0}\right)\right] . \tag{5.19}$$

The corresponding stream function is

$$\varphi(x,y) = AKy - \frac{\gamma}{2\pi}\left[\arctan\left(\frac{y}{x+x_0}\right) - \arctan\left(\frac{y}{x-x_0}\right)\right] . \tag{5.20}$$

The orientation of the dipole with respect to the regional flow allows for quite different operations (Figure 5.6). An orientation with the extraction well upstream and the recharge well downstream would be most appropriate for treating water from a diffuse contamination source that is located way upstream within the capture zone of the extraction well. The converse orientation would be most useful for treating water from a contamination source located between the two wells. Notice that there exist closed streamlines in the second case which ideally would encircle the source region.

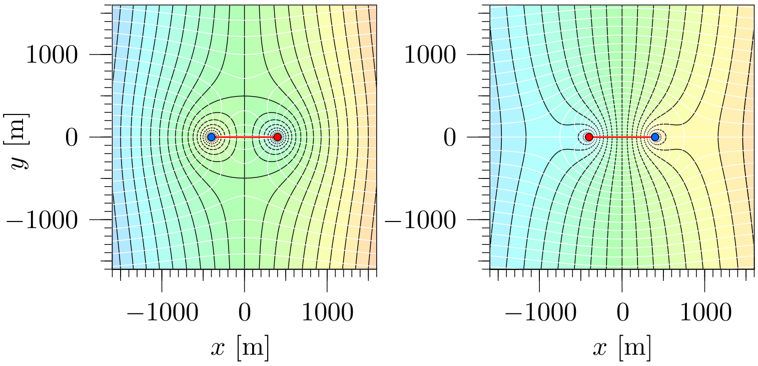


Figure 5.6. Isobars (black lines) and stream functions (white lines) for two configurations of dipole pumping in a uniform regional flow from left to right with extraction (blue dot) upstream and recharge (red dot) downstream or vice versa. Notice that in contrast to Figure 5.5, color shading now represents pressure. The red line again indicates a jump of the stream function by γ. Parameters are identical to those for Figure 5.5.

5.3
Stationary Flow in Heterogeneous Aquifer

As discussed in Section 3.1.2, geologic formations are typically heterogeneous at many scales. As a consequence, all macroscopic material properties, in particular S, θ, and K, become functions of space. We only consider the most simple case here and assume stationary flow, incompressible media, and negligible extraction. Instead of the hydraulic potential ψ_w, we will use the hydraulic head h_w defined by (3.20). This turns Darcy's law into $\mathbf{j}_w = -\mathsf{K}^*(\mathbf{x})\nabla h_w$, where

$$\mathsf{K}^* = \rho g \mathsf{K} \tag{5.21}$$

with dimension $[\mathrm{LT}^{-1}]$ is still called the hydraulic conductivity. Indeed, we will drop the superscript in favor of a more parsimonious notation. The dimension of K is always clear from the context anyway. Notice, however, that the numerical values of K and K^* differ by some four orders of magnitude. With this convention, (5.5) becomes

$$\nabla \cdot \left[\mathsf{K}(\mathbf{x})\nabla h_w\right] = 0 \; . \tag{5.22}$$

We emphasize that (5.22) implies boundary conditions that vary slowly with respect to the relaxation time (5.4).

With the exception of some special cases, (5.22) cannot be solved exactly and we have to get by either with numerical solutions or with analytical approximations. Since analytical approximations are based on rather strong assumptions like macroscopic uniformity and finite correlation length of $\mathsf{K}(\mathbf{x})$ we will generally employ numerical solvers in the following.

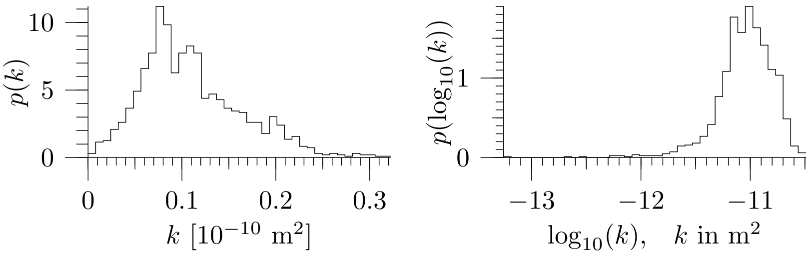


Figure 5.7. Histograms of permeability k (**left**) and of $\log_{10}(k)$ (**right**) at the Borden site [*Woodbury and Sudicky* 1991]. The range is chosen from the minimal to the maximal value measured. (Data courtesy of E. A. Sudicky)

5.3.1
A Field Study: The Borden Site

There are only a few sites for which a sufficient number of high quality measurements are available such that a faithful two- or even three-dimensional representation of the hydraulic structure can be constructed. The reason for this is the high cost for the individual measurement and the "curse of dimensions": 10^3 measurements, for instance, provide an excellent description for a one-dimensional structure, but they are quite insufficient for a three-dimensional representation. One of the sites with detailed information is at the Canadian Forces base near Borden, Ontario. The sandy aquifer underlying it is today considered as a rather uniform formation. It was used as the test bed for many experimental, numerical, and theoretical approaches to subsurface flow and transport.

Hydraulic Structure The hydraulic conductivity field was sampled with a total of 32 cores that were extracted from two orthogonal vertical transects from the depth interval of about 2.5…4.5 m below ground. The cores were separated by 1 m in the horizontal and after extraction were cut into 0.05 m slices. Theses samples were dried and repacked to a porosity of 0.34 before the hydraulic conductivity was measured using a falling head procedure (see Exercise 5.1 for this method). The histogram of permeability k calculated from a total of 1188 conductivity measurements from two orthogonal vertical transects is shown in Figure 5.7 together with the histogram of $\log_{10}(k)$. The distribution is rather broad and in particular includes some very low values as is apparent from $\log_{10}(k)$. They correspond to thin lenses of silty fine-grained sand. The spatial distribution of k in the two transects is represented in Figure 5.8. A geostatistical analysis of $\log(k)$ under the premise of stationarity yields for the correlation lengths in transect AA' 5.1 m horizontally and 0.21 m vertically. The corresponding values for transect BB' are 8.3 m and 0.34 m [*Woodbury and Sudicky* 1991]. The differences between the two transects, which are statistically significant, illustrate the

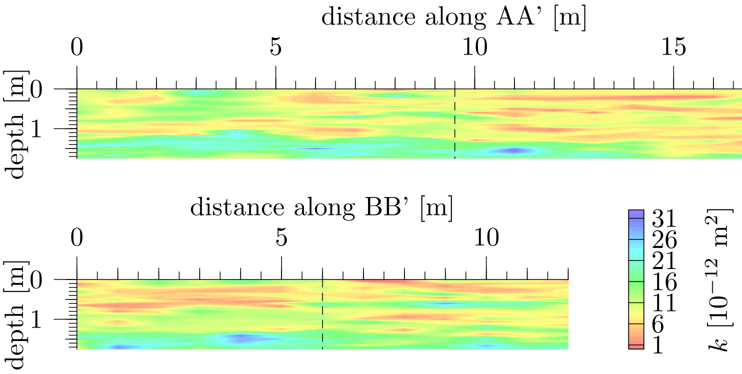


Figure 5.8. Spatial distribution of permeability k in two orthogonal vertical transects at the Borden site [*Sudicky* 1986]. The dashed lines indicate the intersection of the transects. The histogram of the distribution is shown in Figure 5.7. (Data courtesy of E. A. Sudicky)

difficulty in obtaining a reliable three-dimensional representation of the subsurface.

We notice that the measured fields are not stationary. Transect AA' has a considerably higher permeability in the left half than in the right one and there is a region with particularly high permeability in the lower left part. Looking at transect BB', one may hypothesize that this high-permeability zone corresponds to an extended layer that dips gently in AA'-direction. This site serves as a further example for the hierarchical heterogeneity of geologic formations and indicates that statistical properties like the correlation lengths given above are of local significance, at best.

Stationary Flow The measured permeability field is an essential prerequisite for solving (5.22) for the hydraulic head h_w. With Darcy's law, this in turn yields the flow field $\mathbf{j}_w$. Attempting to get such a solution, we encounter three major obstacles: (i) the permeability field is only known in two transects, not in three dimensions, (ii) the boundary conditions are not known, and (iii) only one component of the permeability tensor has been measured. Such limitations are typical for most real applications and demand sometimes severe modeling assumptions which hopefully preserve the character of the solution but almost certainly destroys quantitative details.

To get an impression of the flow field, we assume two-dimensional stationary flow in the vertical transect AA' (Figure 5.8). The flow is driven by a constant horizontal hydraulic gradient and the region is assumed to be bounded by impermeable layers above and below. Finally, permeability is assumed to be isotropic. Figure 5.9 shows the result of the numerical simulation of this model.

As expected, the flow field is not uniform: The high-permeability region in the lower part of the flow domain carries a larger fraction of the flow. In

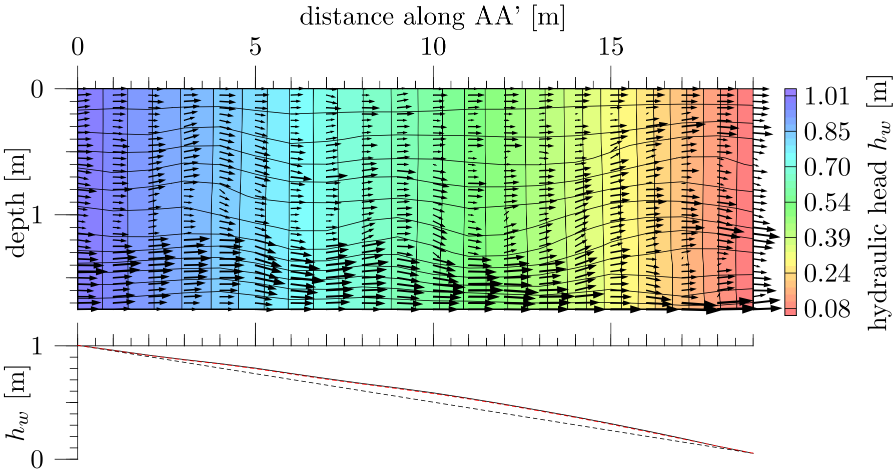


Figure 5.9. Simulated two-dimensional flow at the Borden site assuming impermeable upper and lower boundaries and a constant regional hydraulic gradient leading to a head difference between the left and right boundary of 1 m. The colors and the nearly vertical contour lines represent the hydraulic head. The flow field is indicated by arrows whose length and thickness are proportional to j_w at their tail point and by a few stream lines. Notice that the vertical axis is exaggerated, hence streamlines and lines of constant potential do not appear to be orthogonal even though they are in reality.

contrast, the hydraulic head is rather uniform vertically and is dominated by the horizontal gradient. This result nicely illustrates the previously discussed approximation of horizontal flow in large aquifers with hydraulic conductivity replaced by its vertical integral, the transmissivity T. Such an approximation corresponds to vertical isoplanes of h_w with the mean conductivity between two sufficiently near planes equal to the arithmetic mean of K. In the lower part of Figure 5.9, the vertically averaged hydraulic head of the real transect (black solid curve) is plotted together with the hydraulic head of the corresponding vertically uniform medium (red dashed line). The two curves are hardly distinguishable, thus reinforcing the hydraulic equivalence of the heterogeneous and of the corresponding vertically uniform transect.

Finally, we compare the actual hydraulic gradient with the regional one, which is represented by the dashed black line in the lower part of Figure 5.9. As expected, the higher permeability of the left half of the transect leads to an average hydraulic gradient that is smaller than the regional one and conversely in the right half.

We emphasize again that due to the severe assumptions concerning the hydraulic structure (two-dimensional) and the boundary conditions, the flow field obtained in this simulation is only a rough model of the true flow field at the Borden site.

5.4

Simulated Single-Scale Media

To circumvent the experimental limitations, we consider simulated hydraulic structures which may be of almost arbitrary size and resolution. Specifically, we consider stationary flow through a thin horizontal aquifer of constant thickness whose conductivity field is stationary, isotropic, and characterized by a single scale. In order to minimize the effort for calculating and analyzing the flow field, we are content with a two-dimensional representation. Our aim is to study, for a given spatial structure, the impact of an increasing magnitude of the hydraulic variability.

5.4.1

Hydraulic Structure

As a guide to the construction of the medium, we imagine an unconsolidated sandy formation where the *relative* grain size distribution is the same at everywhere but where the *mean* grain size r varies in space isotropically and with a characteristic scale. Notice that a consequence of this is a constant porosity. However, since r is a characteristic scale for the pore space, the hydraulic conductivity varies proportionally to r^2. Such a model might be considered as the most simple representation of an aquifer that originated from some large flood plain.

We represent $K(\mathbf{x})$ as a realization of some weakly stationary random function which is characterized by its probability distribution and by its autocovariance function. We choose $K(\mathbf{x})$ to be distributed lognormally, hence $Y = \log(K)$ is distributed normally with mean μ_Y and variance σ_Y^2, and the autocovariance function to be isotropic and Gaussian with correlation length ℓ_Y. The magnitude of the spatial variability is thus given by the variance σ_Y^2. A realization with the prescribed spatial properties is generated with the Fourier method based on the Wiener-Khinchin theorem (see Appendix on page 301*f*). It is then transformed linearly such that it has the desired mean and variance. An issue that arises immediately concerns the mean which could be chosen such that the arithmetic mean of K or of $\log(K)$ remains constant as the variance changes, or, alternatively, that the mean flow through the medium remains invariant. We chose $\log_{10}(K)$ to remain constant, which is easy to guarantee. Three fields were then generated such the variance of $\log_{10}(K)$ increases by a factor of 5. The resulting distribution of $\log_{10}(K)$ is shown in Figure 5.10. Notice that already in the field with the smallest variance, the range of K is about an order of magnitude larger than in the AA'-transect at the Borden site. In the field with the largest variance, K covers a range of 14 orders of magnitude. Some further parameters are given in Table 5.1.

The realization generated originally was a periodic square with side length 204.7 m, 2048 nodes with a distance of 0.1 m. For the simulations, a region

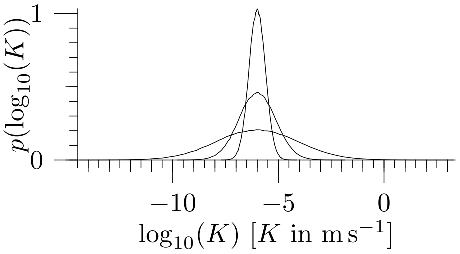


Figure 5.10. Probability distribution functions of $\log_{10}(K)$ used for the simulations. Some parameters of the distributions are given in Table 5.1.

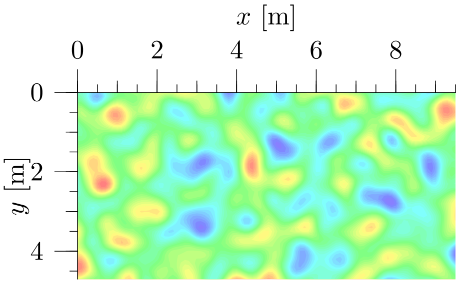


Figure 5.11. Small section from the isotropic log-conductivity field. Correlation length is 1 m. The spatial structure remains the same for all simulations, but the variance of $\log_{10}(K)$ changes.

of 102.3 m×51.1 m was cut from the center in order to avoid artifacts from periodic structures. A small section of such a field is shown in Figure 5.11. As expected, $\log(K(\mathbf{x}))$ is very smooth since it is a realization of a Gaussian random function which are almost always infinitely continuously differentiable. Further notice that for this type of random field, extreme values (red and blue) are always isolated whereas average values (green) form continuous regions. This may be a reasonable structure for some natural formations like a sand bedding, for instance. Others, most prominently crack networks, will show different characteristics, however.

Table 5.1. Some parameters of the hydraulic conductivity distributions used in the simulations. K is given in m s^{-1}. For comparison, the variance of $\log_{10}(K)$ in the Borden AA'-transect is 0.07.

parameter	field 1	field 2	field 3
correlation length ℓ_Y [m]		1.0	
arithmetic mean of $\log_{10}(K)$		-4	
variance of $\log_{10}(K)$	0.15	0.75	3.77
arithmetic mean of K	$1.4 \cdot 10^{-4}$	$6.3 \cdot 10^{-4}$	$3.4 \cdot 10^{-1}$
geometric mean of K	$6.3 \cdot 10^{-4}$	$1.2 \cdot 10^{-5}$	$2.3 \cdot 10^{-8}$
lowest value of K	$4.4 \cdot 10^{-6}$	$9.6 \cdot 10^{-8}$	$1.8 \cdot 10^{-11}$
highest value of K	$2.6 \cdot 10^{-3}$	$1.5 \cdot 10^{-1}$	$1.2 \cdot 10^{+3}$

5.4.2
Simulation of Stationary Flow

We assume the flow to be driven from east to west by a constant average hydraulic gradient. As already found in previous sections, prescribing the exact boundary conditions for a heterogeneous structure is difficult unless they are implied by the setting of the flow domain, i.e., by real impermeable layers which naturally lead to no-flow boundaries or by constant level hydraulic reservoirs like rivers and lakes which lead to constant head boundaries. For numerical studies, on the other hand, we are free to choose the boundary conditions with the sole condition that they do not significantly influence the flow features of interest. Two avenues are typically chosen: (i) the flow domain is fully periodic, which essentially removes all boundaries but may introduce unwanted symmetries, or (ii) a sufficiently small section is cut from a larger domain such that details of the boundaries are not significant anymore. In the following, we choose the second avenue, prescribe constant hydraulic heads across the east and west boundaries to obtain an average hydraulic gradient of about 0.01, and prescribe the north and south boundaries as impermeable. Finally, we cut a 51.1 m×25.5 m section from the center of the simulated domain for further inspection.

Figure 5.12 shows the hydraulic heads and the flow fields for the three conductivity fields characterized in Table 5.1. We notice, as expected, that with increasing magnitude of the conductivity field, (i) the hydraulic head begins to deviate from its mean value which is a plane surface defined by the average hydraulic gradient and (ii) the range of local fluxes increases strongly. More interestingly, we find that with increasing variability the flow becomes ever more localized in a few stream channels. While water fluxes higher than average occur in some 42% of the area in the field with the lowest variability, this reduces to some 23% in the high-variability field (Table 5.2). Correspondingly, the fluxes in these regions are relatively higher: some 5% of the area in the high-variability field carry a water flux that is by more than a factor of 5 larger than the average while in the low-variability field, such high fluxes do not exist at all. At the other end, in some 17% of the area in the high-variability field, water fluxes are more than two orders of

Table 5.2. Some characteristics of the flow fields shown in Figure 5.12. The effective hydraulic conductivity K_{eff} is obtained from (5.24). The spatially averaged flux is $\overline{j_w}$.

parameter	field 1	field 2	field 3
area fraction with $j_w/\overline{j_w} > 1$	0.42	0.34	0.23
area fraction with $j_w/\overline{j_w} < 10^{-2}$	0.00	0.00	0.17
area fraction with $j_w/\overline{j_w} > 5$	0.00	0.01	0.05
$\log_{10}(K_{\text{eff}})$	−4.00	−3.99	−3.92

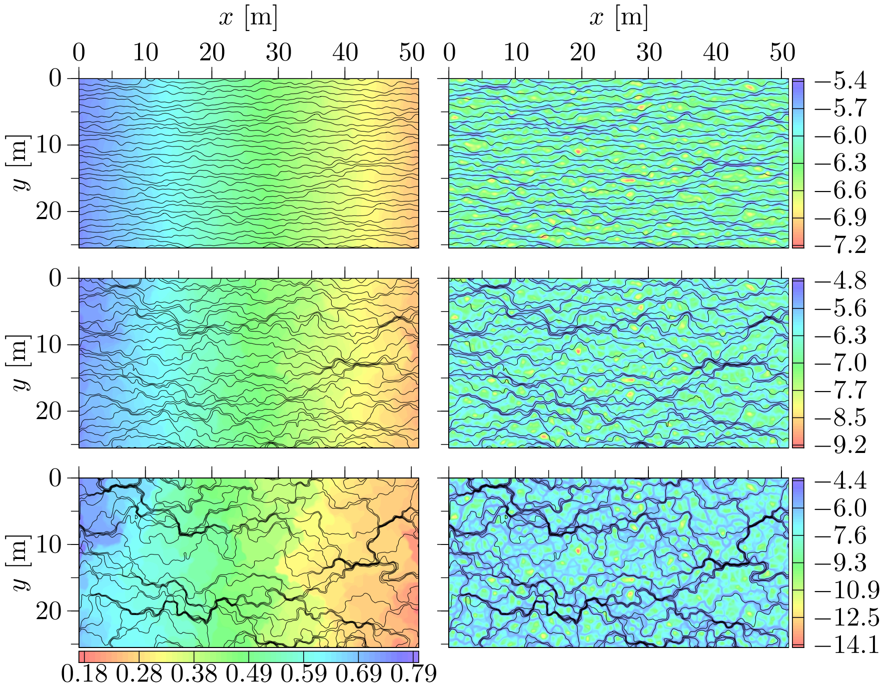


Figure 5.12. Flow in macroscopically uniform media with different degrees of small-scale heterogeneity expressed by the variance of $\log_{10}(K)$ with is 0.15 (**top**), 0.75 (**middle**), and 3.77 (**bottom**) for K given in $\mathrm{m\,s^{-1}}$. Some further parameters of the conductivity fields are given in Table 5.1. The **left** column shows the hydraulic head h_w [m] and the **right** column shows $\log_{10}(j_w)$ with j_w [$\mathrm{m\,s^{-1}}$] the absolute value of the water flux. Equidistant stream lines are shown in black. Notice that the color scale is the same for the hydraulic head in all three cases, but that it is different for the flux fields.

magnitude lower than the average, hence some 17% of the water can safely be denoted as immobile. Again, such immobile water does not exist in the low-variability field. Of particular interest is the spatial arrangement of low- and high-flux zones. As Figure 5.13 shows, low-flux regions tend to form isolated and rather isotropic patches. In contrast, high-flux regions occur as narrow, elongated channels roughly oriented along the mean flow direction. As we will find later, this has important implications for the transport of dissolved substances.

5.4.3
Effective Hydraulic Conductivity

Often, the details of the flow field are not of interest and an approximate coarse-grained description suffices. Since the dynamics of the water flow is

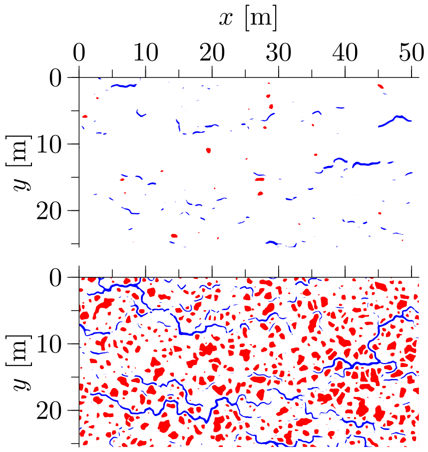


Figure 5.13. High-flux regions with $j_w > 5\overline{j_w}$ (blue) and low-flux regions with $j_w < 10^{-2}\overline{j_w}$ (red) for the simulations with $\mathrm{var}(\log_{10}(K))$ of 0.75 (**top**) and 3.77 (**bottom**). The flow fields are shown in Figure 5.12. There exist no high- or low-flux regions with the given bounds for the field with $\mathrm{var}(\log_{10}(K)) = 0.15$.

linear and flow is assumed to be stationary, we expect to recover Darcy's law also at the correspondingly larger scale. Recall (5.4), however, which shows that stationarity becomes an ever more restricting requirement because equilibration time increases quadratically with scale.

At the new macroscopic scale, which is large compared to the scale of the microscopic heterogeneities, we write Darcy's law as

$$\mathbf{j}_w = -\mathsf{K}_{\mathrm{eff}} \nabla h_w \ , \tag{5.23}$$

where $\mathsf{K}_{\mathrm{eff}}$ is a new material property, called the *effective hydraulic conductivity*. Recall that "macroscopic" and "microscopic" do not refer to a particular size but rather indicate the scale of our current interest and a scale that we wish to absorb into appropriate material properties, respectively. Typical choices for soil physics are 0.1 m for the microscale, the size of typical soil samples to be analyzed in the lab, and 10…1'000 m for the macroscale. Clearly, $\mathsf{K}_{\mathrm{eff}}$ will depend on the microscopic conductivity K, which in turn depends on the fluid's viscosity and on the form of the pore space as represented in the permeability. In addition, $\mathsf{K}_{\mathrm{eff}}$ will also depend on the spatial structure of $\mathsf{K}(\mathbf{x})$. A large effort went into the definition of $\mathsf{K}_{\mathrm{eff}}$ and into its calculation from $\mathsf{K}(\mathbf{x})$ [*Renard and de Marsily* 1997]. In the following, we will illuminate only a few issues that are pertinent to our numerical example.

Coming back to (5.23), we notice that $\mathbf{j}_w$ and h_w are macroscopic quantities different from the microscopic quantities $\mathbf{j}_w^\mu$ and h_w^μ represented for instance in Figure 5.12. It is natural to define $\mathbf{j}_w = \overline{\mathbf{j}_w^\mu}$ and $h_w = \overline{h_w^\mu}$ as spatial averages over the respective microscopic quantities. This allows us to calculate the effective hydraulic conductivity for the flow fields shown in Figure 5.12. To simplify matters, we recall that both, the microscopic conductivity and its spatial distribution were assumed to be isotropic. Hence, we also expect K_{eff} to be isotropic and define it as

$$K_{\text{eff}} := -\frac{j_{w_x}}{\mathrm{d}_x h_w} \; , \tag{5.24}$$

where $\mathrm{d}_x h_w$ is the macroscopic hydraulic gradient that is given by the boundary conditions. The simulations of the three conductivity fields all yielded a value for K_{eff} of about 10^{-4} m s^{-1} (Table 5.2).

Interestingly, we find $\log_{10}(K_{\text{eff}}) \approx \overline{\log_{10}(K(\mathbf{x}))}$, which is a rather simple rule for calculating the effective hydraulic conductivity from the microscopic conductivity field $K(\mathbf{x})$. Indeed, this relation is not accidental. Expanding the divergence in (5.22) and dividing by K yields

$$\nabla^2 h_w^\mu + \nabla Y^\mu \cdot \nabla h_w^\mu = 0 \; , \tag{5.25}$$

where $Y^\mu = \log(K)$ and the superscript μ indicates microscopic quantities. Assuming ∇Y^μ and ∇h_w^μ to be uncorrelated, spatial averaging of (5.25) leads to

$$\overline{\nabla^2 h_w^\mu} + \overline{\nabla Y^\mu} \cdot \overline{\nabla h_w^\mu} = 0 \; . \tag{5.26}$$

If the spatial averaging is done over an REV with constant shape, averaging and differentiation may be exchanged which finally produces

$$\nabla^2 h_w + \nabla \overline{Y^\mu} \cdot \nabla h_w = 0 \; , \tag{5.27}$$

where h_w is now a macroscopic quantity. This shows that two macroscopically uniform fields will have the same effective hydraulic properties provided that the arithmetic mean of their (microscopic) $\log(K)$-fields are identical. This implies that (i) with increasing heterogeneity, i.e., with increasing variance of K, the mean hydraulic conductivity $\overline{K}$ must increase in order to retain the same effective properties and (ii) the autocorrelation function is immaterial for the effective hydraulic properties.

The crucial assumption in deducing (5.27) is that ∇Y^μ and ∇h_w^μ are uncorrelated. However, our result

$$K_{\text{eff}} = \exp\big(\overline{\log(K(\mathbf{x}))}\big) \tag{5.28}$$

can also be deduced rigorously for the case of macroscopically uniform flow in an isotropic, two-dimensional, and lognormal conductivity field with finite correlation length [*Renard and de Marsily* 1997, eq. (33)].

Indeed, (5.28) is one of the few exact results from studying effective properties. Most other results are either conjectures or low-order approximations, typically low-order in $\sigma^2_{\log(K)}$. The most important conjecture, in hydrology referred to as *Matheron's conjecture*, states that (5.28) is a special case of the d-dimensional relation

$$K_{\text{eff}} = \exp\big(\overline{\log(K)}\big) \exp\Big(\sigma^2_{\log(K)} \Big[\frac{1}{2} - \frac{1}{d}\Big]\Big) \; . \tag{5.29}$$

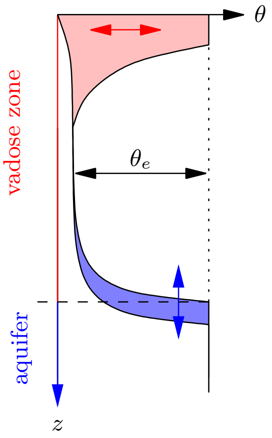


Figure 5.14.
Range of water content in vadose zone (red) and unconfined aquifer (blue). In this example, the surface zone is decoupled from the aquifer. Water content in the surface zone changes according to rainfall and evapotranspiration (see Chapter 6). The groundwater table varies with the regional flow. The water content between the two regions is determined by the hydraulic conductivity and by the mean groundwater recharge rate.

This has been found to be valid in a number of empirical, mostly numerical, studies. There exist a number of further conjectures that in particular account for anisotropy of the conductivity field. However, even for the case of macroscopically uniform flow in a stationary conductivity field with a finite scale, there currently exists no unified method to calculate the effective hydraulic conductivity from a few characterizing quantities of the microscopic conductivity field. The situation becomes even more complicated for non-uniform flow, most importantly near pumping wells, and for hierarchical media where the scale of the microscopic conductivity field is no more small compared to the macroscopic scale. In such cases, the detailed representation of the conductivity field and the numerical solution of the flow equations is required.

5.5
Dynamics of Flow in Unconfined Aquifer

In general, an unconfined aquifer is strongly coupled with the vadose zone and the two cannot be separated. Comprehensive understanding of such a coupled system thus has to await Chapter 6. However, for the special case of a very deep vadose zone the unconfined aquifer may be considered as an entity of its own. This situation is illustrated in Figure 5.14. The thickness required for a vadose zone to qualify as "deep" depends on its soil hydraulic properties and on the statistics of the atmospheric forcing. This will only become clearer in Section 6.3.5.

We first notice that compressibility can be neglected in an unconfined aquifer as it is typically rather shallow with correspondingly low pressures. Hence, the flow can be described with the stationary approximation (5.5). However, the situation is much more complicated than for a confined aquifer

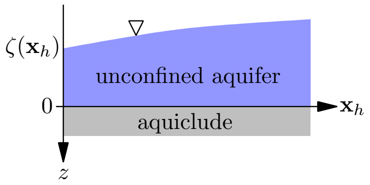


Figure 5.15.
Definition sketch for flow in unconfined aquifer with groundwater table $\zeta(\mathbf{x}_h)$. Notice that here $\mathbf{x}_h$ is a horizontal vector.

because now the upper boundary is moving since the water table itself adjusts to a changing pressure, not just the piezometric head.

We approach the problem in its most simple form and consider a shallow unconfined aquifer in a horizontally extended uniform geologic formation that is bound by a horizontal aquiclude at depth $z = 0$. The hydraulic conductivity K shall be such that one of its major axes is parallel to the vertical axis. We denote the depth of the water table by $\zeta(\mathbf{x}_h) < 0$, with the vertical axis pointing downwards (Figure 5.15).

The conservation of volume may be written as $\partial_t\theta + \nabla \cdot \mathbf{j} = -\gamma$, where γ is the volume extraction rate. We separate all quantities of interest into their horizontal and vertical components, hence $\mathbf{x} = (\mathbf{x}_h, z)$, $\mathbf{j} = (\mathbf{j}_h, j_z)$, and $\nabla = (\nabla_h, \partial_z)$. Integrating over the depth of the aquifer then yields for the conservation of volume

$$\partial_t \int_{\zeta(\mathbf{x}_h,t)}^{0} \theta_e \,\mathrm{d}z + \nabla_h \cdot \underbrace{\int_{\zeta(\mathbf{x}_h,t)}^{0} \mathbf{j}_h \,\mathrm{d}z}_{=:\mathbf{q}} = \underbrace{j_z(\zeta(\mathbf{x}_h,t)) - \int_{\zeta(\mathbf{x}_h,t)}^{0} \gamma \,\mathrm{d}z}_{=:-\Gamma} \;, \tag{5.30}$$

where θ_e in the first term accounts for the fact that the pore space above the groundwater table is not completely empty (see Figure 5.14) and where the flux through the upper boundary, the groundwater recharge from the vadose zone, is summarized into the extraction rate. Notice that the bound water above the water table remains constant in time and thus need not be included in (5.30). Since the formation is uniform, θ_e thus constant, we obtain

$$-\theta_e\partial_t\zeta + \nabla_h \cdot \mathbf{q} = -\Gamma \;, \tag{5.31}$$

where the arguments of ζ have been dropped for clarity.

Next, we invoke Darcy's law for the horizontal flow $\mathbf{q}$. Since a major axis of K is parallel to z, K may also be separated into a horizontal part K_h and the vertical component K_z. The horizontal flux then becomes $\mathbf{j}_h = -\mathsf{K}_h\nabla_h p$, hence

$$\mathbf{q} = -\int_{\zeta}^{0} \mathsf{K}_h \nabla_h p \,\mathrm{d}z \;. \tag{5.32}$$

For a uniform medium, K_h is constant. We further assume that the vertical velocity is negligible compared to the horizontal one, hence assume hydrostatic equilibrium. This is the *Dupuit* assumption. Then, (5.32) yields

$$\begin{aligned}\mathbf{q} &= -\mathsf{K}_h \int_\zeta^0 \nabla_h p \,\mathrm{d}z = -\rho g \mathsf{K}_h \int_\zeta^0 \nabla_h [z - \zeta(\mathbf{x}_h, t)] \,\mathrm{d}z = \rho g \mathsf{K}_h \zeta \nabla_h \zeta \\ &= \frac{1}{2} \rho g \mathsf{K}_h \nabla_h \zeta^2 \;. \end{aligned} \tag{5.33}$$

Inserting this in (5.31), we finally obtain for the dynamics of water flow in an unconfined aquifer

$$-\theta_e \partial_t \zeta + \frac{\rho g}{2} \nabla_h \cdot \left[\mathsf{K}_h \nabla_h \zeta^2 \right] = -\Gamma \;. \tag{5.34}$$

We notice that in contrast to the groundwater equation (5.1), water flow now obeys a nonlinear equation. A very similar situation will be encountered for flow in the vadose zone except that there the nonlinearity is very much stronger.

We comment that sometimes the second to the last term in (5.33) is written with the horizontal transmissivity as $\rho g \mathsf{T}_h \nabla_h \zeta$ and further

$$-\theta_e \partial_t \zeta + \rho g \nabla_h \cdot \left[\mathsf{T}_h \nabla_h \zeta \right] = -\Gamma \tag{5.35}$$

instead of (5.34). Obviously, this formulation is deceiving: the dynamics appears to be linear while in fact it is nonlinear because T_h is proportional to ζ.

5.6
Case Study: Groundwater in Rhein-Neckar Region

Up to now, we have considered artificially constructed media or, in the case of the Borden site, rather well-defined situations. Most problems are much more involved, however, because the subsurface architecture is highly irregular, water flow is forced by a multitude of processes, and, most importantly, the underlying database is notoriously thin.

As an illustrative example, we choose the densely populated and heavily industrialized Rhein-Neckar region located in Germany at 49°N 8°E. Dominant consumers of water are the cities of Mannheim and Ludwigshafen in the urban sector and the large BASF plant in the industrial sector. The hydrogeology of the region is characterized by multistory sedimentary aquifers as sketched in Figure 5.1. The topmost of these aquifers contains literally hundreds of wells, both for pumping and for observation. Surface features include the Rhein river, two ancient meanders, and several gravel pit lakes. Hence the natural setting as well as the human impact on the water resources are rather complicated but typical for many other sites.

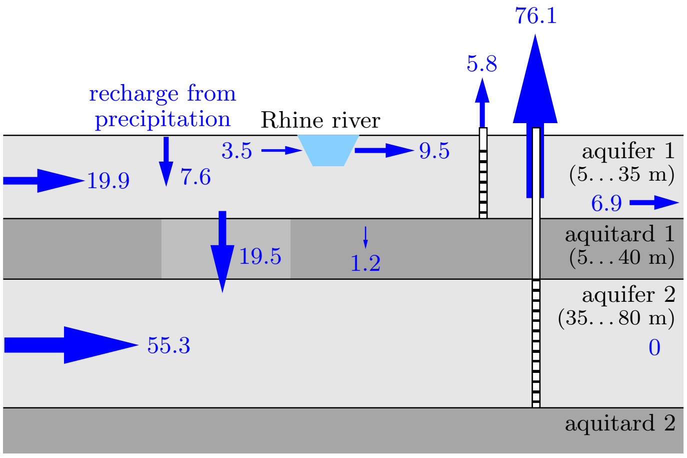


Figure 5.16. Schematic of hydrogeologic architecture and water flow. The aquifers mainly consist of sand and gravel (light gray) while the aquitards are made up of clay and silt (dark gray). There exist a few "windows" in the topmost aquitard (medium gray). They are composed of fine sand and silt and are thus much more permeable. The thickness of the hydrogeologic units varies considerably; the range is given in parenthesis. Notice that the vertical scale is greatly exaggerated: horizontal distances are of order 10 km. The components of the water flows are given in units of 10^3 $\mathrm{m^3 d^{-1}}$. They are aggregated from the numerical simulation which yields spatially resolved flows. [Redrawn after *Wollschläger* 2003]

Wollschläger [2003] studied groundwater flow in a 10 by 16 km subregion in great detail and built a corresponding numerical model. She focused on the two topmost aquifers which are the major suppliers of groundwater. Her findings form the basis for this section.

Prerequisites for a groundwater model are (i) the hydraulic structure of the subsurface, in particular porosity $\phi(\mathbf{x})$ and conductivity $\mathsf{K}(\mathbf{x})$, and (ii) the external forcing of the flow, i.e., boundary conditions and extraction rates. As we will also find for our case, these data are hardly ever available at a resolution and quality required for a faithful groundwater model. Almost invariably, the model must be calibrated, i.e., the parameter fields are adjusted, within the range permitted by the hydrogeologic constraints, such that an optimal agreement between modeled and measured quantities is obtained. Such quantities are typically hydraulic heads at the locations of the observation wells.

5.6.1
Groundwater Model

A rough schematic of pertinent hydrogeologic formations is shown in Figure 5.16. The topmost unconfined aquifer 1 is separated from the deeper

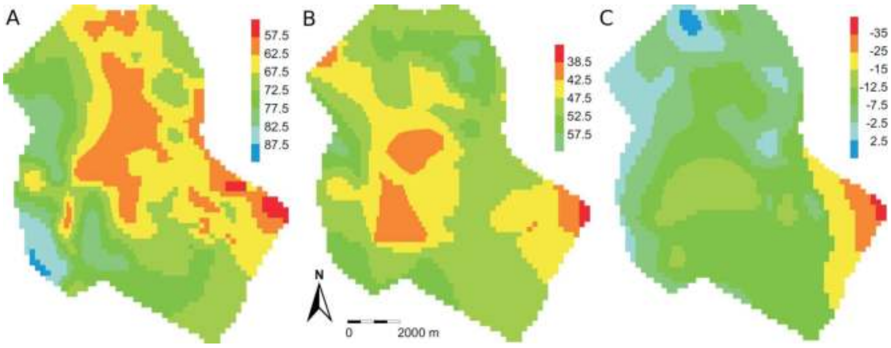


Figure 5.17. Absolute positions of the lower boundaries of (A) aquifer 1, (B) aquitard 1, and (C) aquifer 2 in meters above sea level. [From *Wollschläger* 2003]

aquifer 2 by the leaky aquitard 1, where "leaky" here means that the clay-silt formation is punctured by more permeable sand-silt windows which link the two aquifers. Still the links are sufficiently weak to make aquifer 2 confined.

Water flow in aquifer 1 is driven by groundwater recharge from precipitation, by flow from adjacent regions, by exchange with Rhein river, and by pumping. Anticipating the results of the simulation, the major effect of pumping is actually indirect in that water from aquifer 1 is sucked into aquifer 2 because of the large groundwater abstraction there.

Hydrogeologic Model The subsurface architecture is of course much more complicated than indicated by Figure 5.16. It formed through complicated tectonic and sedimentary processes that started some 45 million years ago. Sedimentation was determined by climate, which changed drastically through the cycle of ice ages, and by the course of the river which meandered throughout the valley. With a high water flow and the river nearby coarse sand and gravel material was sedimented whereas with low water flow and great distance to the river fine silt and clay layers were generated. Ongoing tectonic movements include the sinking of the Graben with $0.2 \ldots 0.9\ \mathrm{mm\,y^{-1}}$ which is the origin of the sedimentation. The Graben sinks somewhat faster in the East. This leads to a corresponding dipping of the originally horizontal sedimentary layers. Figure 5.17 gives a more detailed representation of the subsurface structure. It reveals that the large-scale features were modified by a number of smaller-scale processes like local meandering and flooding. For instance, the bottom of aquifer 1, (A), hints at a broad river valley with insular elements. A still closer view, as it is provided by core drillings, yields even more details of the intricate multiscale structure (Figure 5.18). Apparently, the various aquifers and aquitards are not uniform entities but represent aggregated abstractions. Taking aquitard 1 as an example (Figure 5.19), we indeed find a highly variable thickness that ranges between some 40 m and less than 5 m, and a variable facies composition.

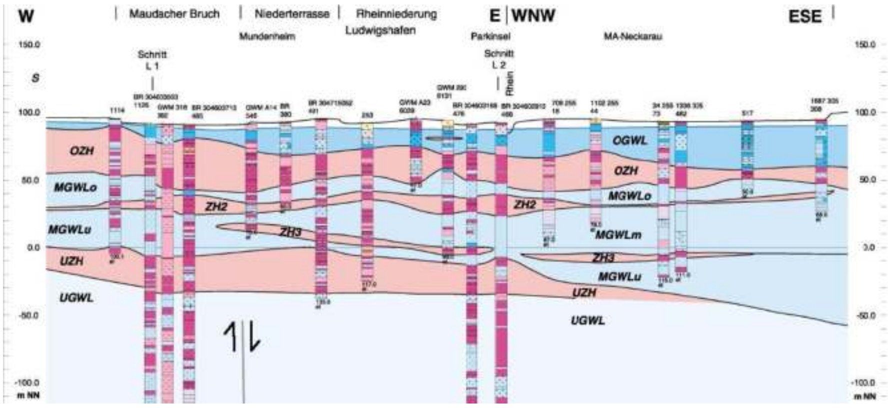


Figure 5.18. Hydrogeologic cross-section through part of Rhein valley located in the lower third of Figure 5.17. The stratigraphy obtained from the core drillings reveals the substructure of the large hydrogeologic units. OGWL corresponds to aquifer 1 in Figure 5.16, OZH to aquitard 1, and MGWL to aquifer 2. [Extracted from cross-section Q3 of *HGK* 1999]

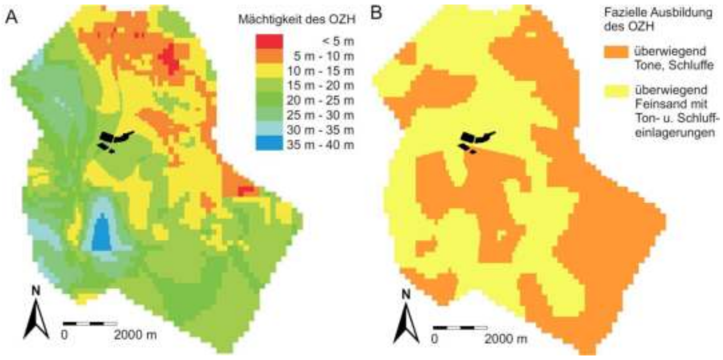


Figure 5.19. Thickness of aquitard 1 (A) and distribution of facies (B). The black areas show some gravel pit lakes. [From *Wollschläger* 2003]

Once the shape of the various hydrogeologic units is known, their hydraulic properties are desired. These depend on the material, specifically its texture and mineral composition, and on its alteration since the time of deposition, in particular its consolidation by the increasing load of overlaying material. Again expected, the hydraulic properties vary significantly within any hydrogeologic unit and estimates based on the facies alone are very imprecise. More accurate values are obtained from pumping tests, more than 500 of which have been performed in the Rhein-Neckar region. Based on their results, maps of storage coefficient S and hydraulic conductivity K are constructed and collected into hydrogeologic maps [*HGK* 1999]. The ranges of these fields for the three topmost hydrogeologic units are given in Table 5.3. Looking at the values, first notice that they are given in units of m^{-1} and $\mathrm{m\,s}^{-1}$, respectively. This corresponds to writing (5.1) in terms of hydraulic head $h = p/[\rho_w g]$ instead of p. Next, we realize that the storage coefficient of aquifer 1 is some 5 orders of magnitude larger than that of aquifer 2. This reflects the difference between an unconfined and a confined aquifer. Finally, the conductivity of the aquitard is some 6 orders of magnitude lower than that of the aquifers, except for the "windows" where the ratio is only some 3 orders of magnitude. Obviously, the location and the exact value for the conductivity of these "windows" is crucial for a faithful description of the coupling between the two aquifers.

External Forcing Water flow is forced by a multitude of processes which are, naturally, most diverse for the topmost aquifer. They include (i) horizontal fluxes across the vertical boundaries from parts of the aquifer beyond the domain of interest, (ii) groundwater recharge from infiltrating precipitation, (iii) exchange with surface waters like rivers and lakes, and (iv) water abstraction wells. Mathematically, these forcings enter the model as boundary conditions for the groundwater equation (5.1) and as extraction term γ, respectively. Hence, their accurate representation is a prerequisite for reliable simulations.

All the forcings are very difficult to quantify, although for quite different reasons: (i) Fluxes across vertical boundaries clearly depend on the dynamics beyond our region of interest and are thus *a priori* unknown. A practical remedy is to measure the hydraulic head along the boundary – fluxes are next to impossible to measure – and to prescribe them as boundary condition in the solution of (5.1). (ii) Precipitation is comparatively simple to measure and data are available at reasonable spatial resolution from weather services. However, a significant fraction of the precipitation does not contribute to groundwater recharge but runs off from the surface, directly evaporates from the soil, or transpires through plants. This fraction depends in a complicated way on many factors including soil type and vegetation, actual weather and its history over the past week to months, surface morphology, and depth of groundwater table. Figure 5.20 shows the estimated mean annual recharge for the Rhein-Neckar region. Low and even negative recharge rates are found

Table 5.3. Ranges for storage coefficient $S^* = \rho_w g S$ and hydraulic conductivity $K^* = \rho_w g K$ of topmost hydrogeologic units in Rhein-Neckar region. In addition, the coefficient D_p of pressure diffusion is given for which we assume that K and S are strongly correlated. [From *HGK* 1999; *Wollschläger* 2003]

	storage coefficient S^* $[\mathrm{m}^{-1}]$	conductivity K^* $[10^{-3}\ \mathrm{m\,s}^{-1}]$	$D_p = K/S$ $[\mathrm{m}^2\mathrm{s}^{-1}]$
aquifer 1	0.09…0.21	0.7…1.4	0.007…0.008
aquitard 1		0.000001…0.001	
aquifer 2	$10^{-6}\ldots10^{-5}$	0.1…0.3	30…100

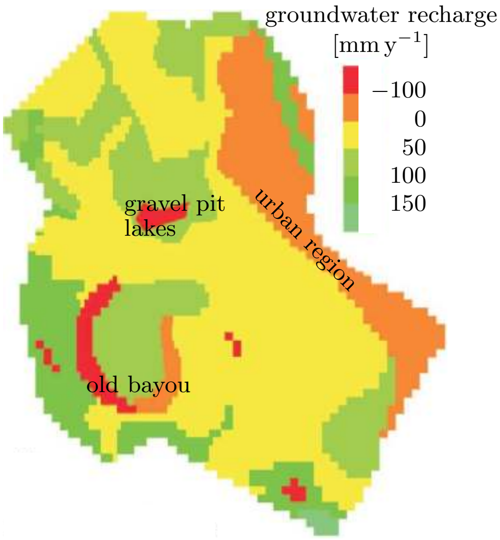


Figure 5.20.
Estimated mean annual groundwater recharge in Rhein-Neckar region. Notice that recharge is very different from precipitation which is rather uniform across this region. [From *Wollschläger* 2003]

in the Eastern, dominantly urban part, where precipitation is collected by the sewer system, in the wetlands of the old Rhein bayou, and around the gravel pit lakes. Positive recharge results from the elevated terraces with a dominantly rural character. (iii) Quantifying the exchange with surface waters may at first appear to be rather simple since the hydraulic conductivity is obtained easily. However, the water bodies are typically separated from the groundwater by a fine-textured sediment layer whose properties are little known, but expected to be highly variable in space. (iv) Also the extraction rate at the various production wells would appear to be well-known. While this is indeed the case, the major difficulty is that these data tend to be considered as sensitive information by the operators, industrial complexes and water producers.

5.6.2
Model Calibration

Ideally, the hydraulic parameter fields $S(\mathbf{x})$ and $\mathsf{K}(\mathbf{x})$ of a groundwater model are specified based on maps which delineate the hydrogeologic units and a set of pumping tests which provide the values of S and K for these units. Obviously, the situation is much more complicated in reality: We already found with Figure 5.18 that such units are not uniform and discussed the multiscale heterogeneity of the subsurface in Chapter 3 above. This is further exacerbated by the high cost of an individual measurement which requires at least one and typically several wells. Consequently, the parameter fields are notoriously ill-defined and simulations based solely on them are typically unable to capture the hydraulic dynamics quantitatively.

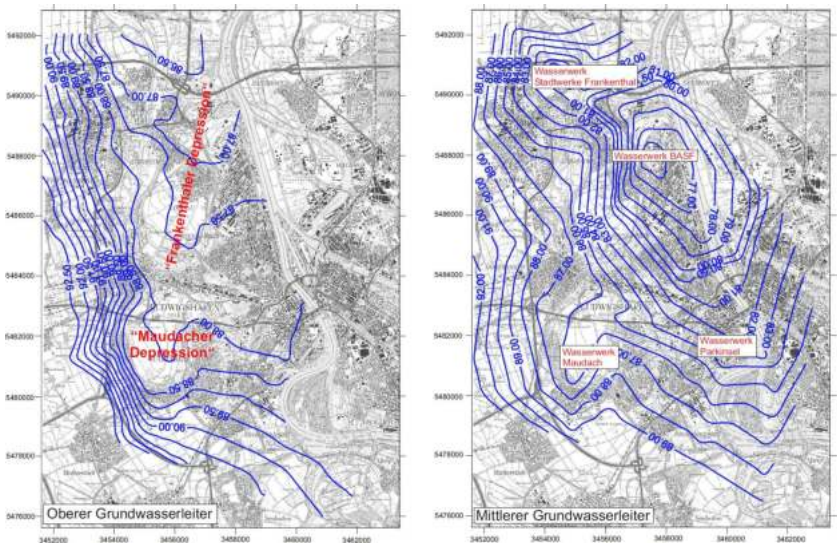


Figure 5.21. Measured hydraulic heads in aquifers 1 (left) and 2 (right). Large-scale depressions result from the strong abstraction of groundwater either directly or, as in the case of the "Maudacher depression", indirectly through the link with another aquifer. [From *Wollschläger* 2003]

In general, there is much more information available than the hydrogeologic maps, however. These are in particular the hydraulic heads which are monitored routinely at a number of observation wells, or the levels of lakes and rivers. Figure 5.21 shows an example of such data: contour lines interpolated from measured hydraulic heads in aquifers 1 and 2 in the Rhein-Neckar region. In order to use such data, a groundwater model must be available in the first place. It will obviously be based on the *a priori* available hydrogeologic data. Elements of such a model for the two topmost aquifers of the Rhein-Neckar region are shown in Figure 5.22. Here, we only look at the case of stationary flow as it is for instance of interest for sustainability studies. Simulating the flow thus entails solving (5.5) with prescribed conductivity and extraction fields $\mathsf{K}(\mathbf{x})$ and $\gamma(\mathbf{x})$, respectively, and to drive it with constant boundary conditions. This was done with a model based on a finite differences discretization (Appendix B).

Calibrating a groundwater model means that $\mathsf{K}(\mathbf{x})$, and possibly $\gamma(\mathbf{x})$, are adjusted such that (i) they stay within the permissible range determined by the hydrogeologic information and by the production data and that (ii) simulated hydraulic heads at the observation points are in optimal agreement with the measurements. A comparison between simulated and measured hydraulic heads after calibration of the model is shown in Figure 5.23.

Calibration is an essential step in the development of a groundwater model. It corresponds to the optimal integration of all the available information

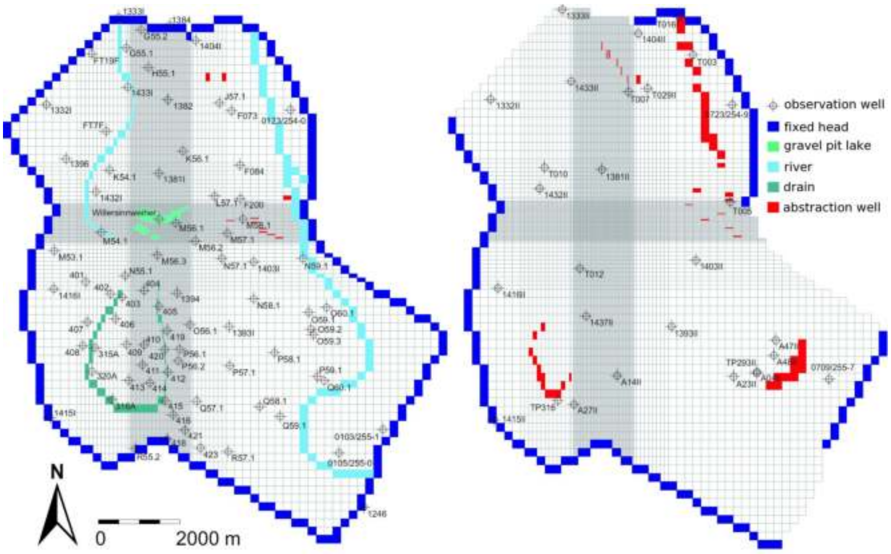


Figure 5.22. Representations of aquifer 1 (left) and 2 (right) in numerical groundwater model. The gray lines indicate the finite differences grid on which the groundwater equation is discretized. This particular model aimed at studying the coupling between gravel pit lakes and groundwater. To this end the grid was refined in the center to increase the resolution there. Boundary conditions for the flow are either a prescribed hydraulic head (blue) or a vanishing flow. Abstraction wells or fields of them are shown as red areas. For aquifer 1, the locations of river courses, lakes, and of the old bayou which acts as a drain are indicated. [From *Wollschläger* 2003]

about the specific region and the flow through it into a model of a prescribed form, here given by the stationary groundwater equation (5.5). In the wider field of geosciences, this integration is referred to as *data assimilation* and there exist several highly sophisticated methods to accomplish it.

Notice that calibration must not be confused with *validation*, which attempts to ascertain that the simulated flow indeed represents the real one: Since all the available information has been used for the calibration, there is none left for the validation. This is a very general difficulty and there exist a number of approaches. These include: (i) *jack-knifing*, exclusion of a subset of the measurements from the calibration set and using it to judge the quality of the model, (ii) χ^2-*statistics*, where the deviation between simulated and measured values (hydraulic heads,...), relative to the measurement uncertainty, quantifies the correctness of the model, and (iii) use of information which is completely independent of the calibration process, for instance the distribution of tracers like artificial solutes from tracer tests, geochemical indicators, isotopes, or temperature. Whatever is done for validation, the information gained there will again be used to improve the model, that is

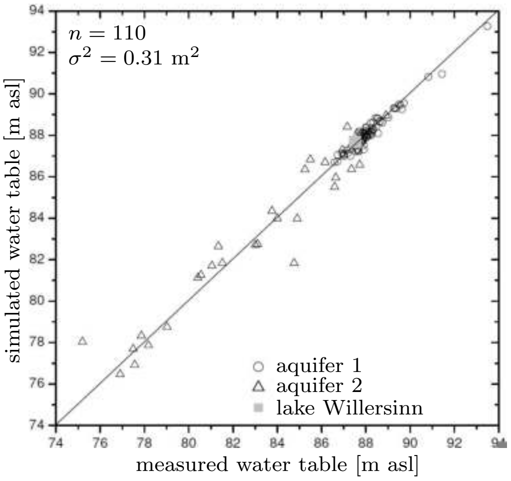


Figure 5.23.
Calibration results of groundwater model for Rhein-Neckar region. A total of 110 observation points were used and the variance σ^2 of the deviation between measured and simulated hydraulic head was 0.31 m^2. [From *Wollschläger* 2003]

to calibrate it further. This leads to a continuous evolution towards an ever more reliable model. Changes between such iteration can actually be quite dramatic. For instance, the conductivities of aquifer 1 given in Table 5.3 are by about a factor of 2 higher than in the previous hydrogeologic map that dates from 1987.

5.6.3
Simulation

Groundwater flow in the Rhein-Neckar region, as calculated with the calibrated groundwater model, is shown in Figure 5.24. The streamlines reveal a quite complicated three-dimensional flow with flow directions in the two aquifers often quite different, in some subregions even opposite. Generally, water flow is from the upper aquifer 1 to the lower aquifer 2 where most of the water is abstracted. This transfer occurs mostly through the hydraulic windows in the intermediate aquitard where the thickness is small and the material contains a high fraction of fine sand. Such conditions prevail in the northern part of the region west of the Rhein river (Figure 5.19). There, the streamlines indicate that water infiltrates from the river into aquifer 1, travels westward to the extended hydraulic window where it sinks into aquifer 2 and travels back east to the gallery of pumping wells.

An aggregated perspective of the groundwater flow in the Rhein-Neckar region was already given as a lookahead in Figure 5.16. It reveals that all the water that enters aquifer 2 is abstracted and that some 27% of this water was transferred across the aquitard. Rhein river is a net source of groundwater. This is mainly so in the northern part of the region whereas in the southern part it is a net sink, as can be gathered from Figure 5.24.

Finally, we notice that the time scale of the water movement ranges between a few decades for infiltration from Rhein river and more than a century

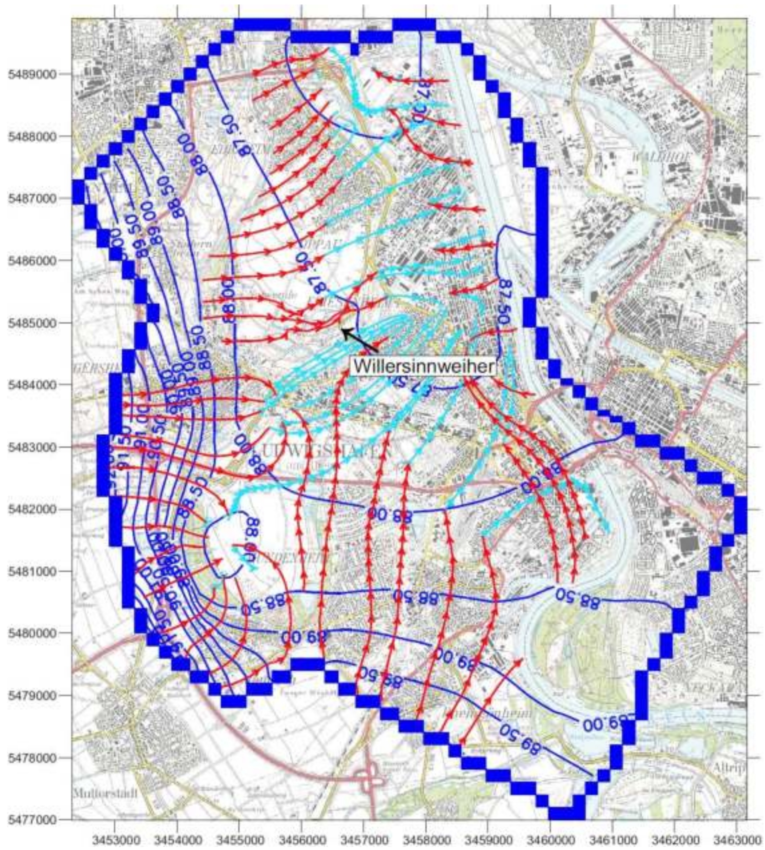


Figure 5.24. Simulation of stationary groundwater flow in Rhein-Neckar region. Blue lines are contours of the hydraulic head in aquifer 1 given as height [m] above sea level. The arrowed lines represent selected streamlines in aquifer 1 (red) and 2 (cyan). The distance between the arrows corresponds to a travel time of 5 years. Notice that all of the streamlines start in aquifer 1. [From *Wollschläger* 2003]

for the water coming from the eastern boundary. From the perspective of water quality, this implies that groundwater that is currently pumped reflects the environmental quality some fifty years back, at least for non-adsorbed chemicals like nitrate and pesticides.

Exercises

5.1 Falling Head Permeameter Consider a vertical, homogeneous, water-saturated soil column of length ℓ with water ponding to height h_0, i.e., the water table is at height h_0 above the upper end of the soil column. At time $t = 0$, the column is allowed to drain freely by removing the blocking at the lower end. Calculate the height of the water table as a function of time, $h(t)$, and deduce the hydraulic conductivity from corresponding measurements. [Such an instrument is called a falling head permeameter.]

5.2 Flow through Layered Column Consider a 1 m long vertical column that is water-saturated with a constant water table at the upper end and free outflow at the lower end. It is composed of uniform sand ($K_w = 10^{-4}\ \mathrm{m\,s^{-1}}$) with a 0.1 m thick layer of uniform silt ($K_w = 10^{-6}\ \mathrm{m\,s^{-1}}$) in the middle. Calculate the pressure $p(z)$ in the water phase in (dynamic) equilibrium and the flux. How do these quantities change if the silt layer is not in the middle?

5.3 Permeability and Characteristic Pore Size Given the data in Table 5.3 on page 125, calculate the permeability of the hydrogeologic units and the corresponding equivalent pore diameters.

5.4 Zone of Disturbance Given the data in Table 5.3 on page 125 for aquifers 1 and 2, calculate the characteristic radius of the zone influenced by pressure fluctuations resulting from groundwater abstraction with a diurnal variation.

6
Soil Water Flow

Water flow in soils, in the vadose zone, is characteristically different from flow in aquifers in that (i) the water is bound, consequently the flow predominantly vertical, (ii) the flow involves variable fractions of the pore space, and (iii) it is strongly coupled with the atmosphere through rainfall and evapotranspiration, which leads to stochastic fluctuations of the water content. As we will find, the depth interval over which these fluctuations are dissipated decreases with increasing frequency. Hence, the time scale of the water flow increases rapidly with distance from the soil surface.

In reality, not only the water content of soils is highly variable but the soils themselves. This results most importantly from shrinking and swelling during drying and wetting cycles or from corresponding changes in the wettability. We will not consider these sometimes dramatic modifications here and assume a rigid and perfectly wettable porous medium. This is a reasonable approximation for all soils at some distance from the ground surface and for soils with low contents of clay and organic matter throughout the profile.

6.1
Dynamics of Soil Water

Water flow in soils is *a priori* a multiphase process since a change in water content induces a complementary change in air content. However, since air is much more mobile than water – $\mu_a/\mu_w = 0.0136$ at 10°C – the water phase decouples from the air phase when the latter becomes continuous and its volume fraction θ_a exceeds a critical value. In this so called *degenerate multiphase regime* the conductivity for the air phase is much higher than that for the water phase (Figure 6.1). Any change of θ_a, induced by a corresponding change of θ_w, is thus easily accommodated by gradients of p_a that are much smaller than those of p_w. The only manifestation of the second phase in this regime is the variable saturation and the ensuing change of the material properties.

The critical air content θ_a^{crit}, where the air and the water phase decouple, may actually be quite high, on the order of 0.1. This is because the residual

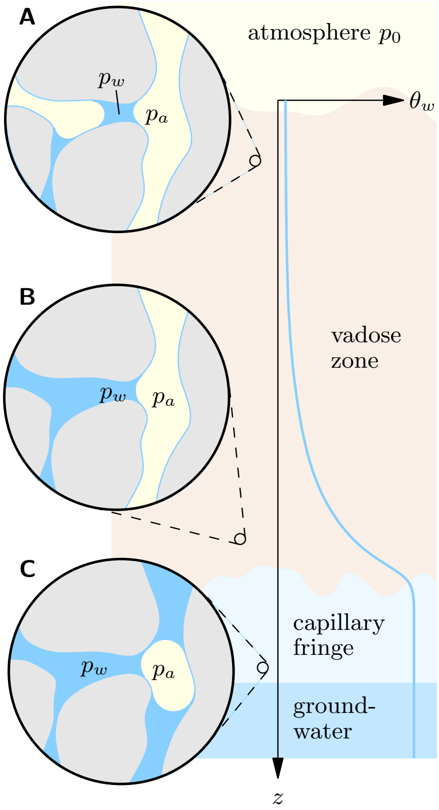


Figure 6.1. Flow regimes in soils. Far from groundwater, water content θ_w is typically small enough that the air phase is continuous through large conduits: gradients of p_a are negligible compared to those of p_w. The two phases are decoupled leading to a *degenerate multiphase regime* (**A**). Nearer to groundwater, or with strong infiltration fronts, θ_w increases such that the air phase remains continuous but air content θ_a is so small that gradients of p_a are no longer negligible. The two phases become strongly coupled in this *continuous multiphase regime* (**B**). With θ_w increasing even further – in the capillary fringe or above intermediate low-conductivity layers – the air phase becomes discontinuous (residual) with air bubbles typically blocking large openings. In this *discontinuous multiphase regime* (**C**), air flow is no more continuous.

air content θ_a^{res} – the air volume that is not connected to the atmosphere – is usually already quite large (see Figure 3.22 on page 60). For $\theta_a^{\text{crit}} > \theta_a > \theta_a^{\text{res}}$, the *continuous multiphase regime* which often occurs in a rather small range of θ_a that is only a few 0.01 wide, air and water phase become strongly coupled and the Buckingham-Darcy flux law (3.59) is no more applicable.

Finally, there is the *discontinuous multiphase regime* with $\theta_a < \theta_a^{\text{res}}$ which always reigns in the capillary fringe but may also occur above and within low-permeability soil layers. In this regime, the conductivity for the air phase is many orders of magnitude smaller than in continuous multiphase regime and essentially determined by the diffusion coefficient of air in water.

We only introduced the notion of a discontinuous multiphase regime for the wet end, where the air phase becomes discontinuous. Of course, there is also a residual water phase at the dry end in the sense that the water phase becomes disconnected except for the omnipresent film on the soil matrix and some pendular rings. However, vapor diffusion is a rather fast transport path, at least when compared to gas diffusion through water as it is relevant for residual air. As a consequence, the *effective* hydraulic conductivity – total water flux in liquid and vapor phase per unit gradient – decreases much more gradually at the dry end than the conductivity for air at the wet end.

6.1.1

Degenerate Multiphase Regime

We consider the air phase as arbitrarily mobile such that the pressure gradient required to move it is negligible compared to the one required for moving the water phase. The pressure p_a in the air phase is thus constant and equal to that of the atmosphere. The water flux is then described by the Buckingham-Darcy law (3.59) which we write as

$$\mathbf{j}_w = -\mathsf{K}_w(\theta_w)\nabla\psi_w = -\mathsf{K}_w(\theta_w)[\nabla\psi_m - \rho_w\mathbf{g}] \ , \tag{6.1}$$

where ψ_w is the *water potential* (3.16) and ψ_m is the *matric potential* (3.19). Recall that ψ_m equals the pressure jump across the water-air interface and is related to the curvature of the interface by the Young-Laplace equation (3.2). Inserting (6.1) into $\partial_t\theta_w + \nabla\cdot\mathbf{j}_w = 0$, the formulation of the conservation of water volume, yields

$$\partial_t\theta_w - \nabla\cdot\big[\mathsf{K}_w(\theta_w)[\nabla\psi_m - \rho_w\mathbf{g}]\big] = 0 \ , \tag{6.2}$$

the *Richards equation* which was presented in a seminal paper by *Richards* [1931]. For a complete formulation, (6.2) must be supplemented with descriptions of the material properties, specifically with the relations $\theta(\psi_m)$ and $\mathsf{K}_w(\theta_w)$ that were introduced in Section 3.4.

Since at least $\theta(\psi_m)$ is a hysteretic relation, we end up with a rather complicated model which is still a major hurdle for theoretical analysis as well as for numerical simulations. The situation is much simpler for processes that evolve on a single branch of $\theta(\psi_m)$. This is for instance the case for pure infiltration and for pure evaporation. It is then convenient to incorporate $\theta(\psi_m)$ directly into (6.2). Obviously, there are two ways to do this.

ψ-Form of Richards Equation Inserting $\theta(\psi_m)$, which is a unique function on an individual branch, directly into (6.2) yields

$$C_w(\psi_m)\partial_t\psi_m - \nabla\cdot\big[\mathsf{K}_w(\psi_m)[\nabla\psi_m - \rho_w\mathbf{g}]\big] = 0 \ , \tag{6.3}$$

where C_w is the *soil water capacity function* introduced in (3.38) and $\mathsf{K}_w(\psi_m)$ is short for $\mathsf{K}_w(\theta(\psi_m))$. It is this form of the Richards equation that is solved by most current numerical codes.

θ-Form of Richards Equation Alternatively, one might invert a branch of $\theta(\psi_m)$. Away from saturation, this is possible in most cases since the individual branches are typically strictly monotonic. Then, a few manipulations lead to

$$\partial_t\theta_w + \underbrace{\mathbf{V}_w(\theta_w)\cdot\nabla\theta_w}_{\text{advection}} - \underbrace{\nabla\cdot\big[\mathsf{D}_w(\theta_w)\nabla\theta_w\big]}_{\text{dispersion}} = 0 \tag{6.4}$$

with

$$\mathbf{V}_w(\theta_w) := \rho_w \frac{\mathrm{d}\mathsf{K}_w}{\mathrm{d}\theta_w}\mathbf{g} \,, \qquad \mathsf{D}_w(\theta_w) := \frac{\mathrm{d}\psi_m}{\mathrm{d}\theta_w}\mathsf{K}_w(\theta_w) \,. \tag{6.5}$$

This may be interpreted as a nonlinear advection-dispersion equation for the water content with advection velocity $\mathbf{V}_w(\theta_w)$ and soil water diffusivity $\mathsf{D}_w(\theta_w)$. Examples for $V_w(\theta_w)$ and $D_w(\theta_w)$ in a one-dimensional soil are shown in Figure 6.17 on page 161.

The formulation (6.4)–(6.5) is attractive for some theoretical analyses because it can be approached with the Boltzmann transform [e.g., *Philip* 1957*a*]. It has a major drawback, however: in heterogeneous media, θ_w is in general discontinuous while ψ_m is continuous. The latter is a direct consequence of the flux law (6.1), since a discontinuity of ψ_m leads to a diverging water flux. Since different materials have different $\theta_w(\psi_m)$-relations, θ_w cannot be continuous in general. Because of this difficulty, the θ-form of the Richards equation is hardly ever used for realistic situations which typically involve heterogeneous formations.

6.1.2
Continuous Multiphase Regime

Consider the simultaneous flow of water and air in the regime where both phases are continuous. This means that any two points in the same phase can be connected by a path within that phase. Further assume for simplicity that also the air phase is incompressible. While this is not entirely justifiable in general, it is reasonable here since we only consider small pressure fluctuations. Finally, to keep the notion simple, we focus on a uniform medium and one-dimensional horizontal flow, thereby neglecting gravity.

The flux of each phase is assumed to be described by the Buckingham-Darcy law (3.59) with the potential defined according to (3.16). Since flow is horizontal the gravity term drops out and the potential for phase i becomes $\psi_i = p_i - p_0$, where p_i is the pressure. We choose the ambient air pressure as reference p_0. The flux laws then become

$$\begin{aligned} j_a &= -K_a(\theta_a)\partial_x\psi_a \\ j_w &= -K_w(\theta_w)[\partial_x\psi_m + \partial_x\psi_a] \,, \end{aligned} \tag{6.6}$$

where ψ_w was written as

$$\psi_w = [p_w - p_a] + [p_a - p_0] = \psi_m + \psi_a \tag{6.7}$$

with the matric potential ψ_m defined in (3.19). In the literature, ψ_m is sometimes called the *static capillary potential* and $\psi_m + \psi_a$ the *dynamic capillary potential*.

Notice the asymmetry in the flux laws (6.6). It stems from the fact that ψ_m is referred to the pressure in the air phase, which may vary in space and time, while ψ_a is referred to p_0. The reason for this is that a path to the reference pressure exists in the air phase, but not in the water phase.

Example: Horizontal Air Flow Consider a thin horizontal slab of uniform porous material with water content $\theta_w < \phi$. In equilibrium, the water will be distributed such that ψ_m is constant everywhere. Now impose a horizontal air flow through this medium by applying the pressure gradient $\partial_x p_a = \partial_x \psi_a$ in the air phase.

Assume that the water phase in the medium is not connected to any outside reservoir. Then, water cannot flow, $j_w = 0$, and (6.6) requires $\partial_x \psi_m = -\partial_x \psi_a$ for equilibrium. Hence a gradient in the matric potential will develop such that $\partial_x \psi_w = 0$ or, with (6.7), $p_w = \text{const}$. Obviously, this is what we expect since water will flow due to a gradient in the (microscopic) pressure. Since ψ_m changes with x also the water content will change according to the soil water characteristic $\theta_w(\psi_m)$: water content will increase towards the outlet. We notice in passing that in reality, the situation is somewhat more involved in that water will evaporate and be transported by the air stream with the evaporation rate determined by the difference between ψ_m and the chemical potential of vapor in the air phase, which is a function of humidity. Adjusting the humidity such that the corresponding potential equals the mean value of ψ_m, water will condense at the air inlet and evaporate at the air outlet thus inducing a small flux in the water phase.

With the flux laws (6.6), we may now formulate the conservation of fluid volumes as

$$\begin{aligned} \partial_t \theta_a - \partial_x [K_a(\theta_a) \partial_x \psi_a] &= 0 \\ \partial_t \theta_w - \partial_x \big[K_w(\theta_w)[\partial_x \psi_m + \partial_x \psi_a]\big] &= 0 \\ \partial_t \theta_a + \partial_t \theta_w &= 0 \ , \end{aligned} \tag{6.8}$$

where the last equation states that there are only two fluids to fill the pore space. These coupled nonlinear partial differential equations have to be solved for describing the simultaneous flow of water and air in soil.

As a first step to better understand the dynamics, we substitute $\partial_x \psi_a$ from (6.6) to arrive at a modification of Buckingham's conjecture, namely

$$j_w = -K_w(\theta_w) \partial_x \psi_m + \frac{K_w(\theta_w)}{K_a(\theta_a)} j_a \ . \tag{6.9}$$

In addition to the flux $-K_w(\theta_w)\partial_x \psi_m$ which corresponds to the degenerate case, (6.1), there appears a contribution from the air flux, weighted by the ratio of conductivities of the two phases. To assess its importance, consider the typical situation in soils where the air flux occurs in response to a transient water flux that changes θ_w. Clearly, the air flux is at most equal to the water flux, typically much smaller, and with opposite sign. We thus look at the ratio K_w/K_a and, recalling (2.38), write it as

$$\frac{K_w(\theta_w)}{K_a(\theta_a)} = \frac{\alpha_w}{\alpha_a} \left[\frac{\ell_w}{\ell_a}\right]^2 \frac{\mu_a}{\mu_w} \frac{\theta_w}{\theta_a} \ . \tag{6.10}$$

Notice that (i) $\alpha_w \ell_w^2 \theta_w$ may be interpreted as the permeability of the water-filled pore space, (ii) the characteristic length ℓ_w of this fraction is a rapidly decreases function of θ_w since the water phase tends to occupy smaller pores with stronger capillary forces, and (iii) the geometry parameter α_w does not depend strongly on θ_w as long as the pore geometry does not change strongly with size. The analogous statements are valid for $\alpha_a \ell_a^2 \theta_a$ as permeability of the air-filled pore space. The two permeabilities, expressed as functions of the respective phase saturations, will have the same value for full saturation, $\theta_w = \phi$ and $\theta_a = \phi$, respectively. Their decrease will be characteristically different, however, with a much stronger initial decrease for the water phase. The reason for this is that water retreats from the largest pores whereas air retreats from the smallest ones.

With the above understanding, we assume for the individual terms of (6.10) that α_w/α_a is of order 1, $[\ell_w/\ell_a]^2$ is typically much smaller than 1, and $\mu_a/\mu_w \approx 0.01$ (0.0136 at 10°C). However, the ratio θ_w/θ_a between the volume fractions becomes arbitrarily large as θ_a decreases. Still, as long as θ_a is larger than a few percent, the ratio K_w/K_a is much smaller than 1 and the modification to Buckingham's conjecture is small. Clearly, near saturation it begins to dominate the water flux and the traditional Buckingham-Darcy law becomes useless.

We notice that the modification to Buckingham's conjecture accounts for the pressure gradient required to move the air phase following a changing water phase. To make this explicit, write $\partial_x j_a = -\partial_t \theta_a = \partial_t \theta_w$ from the conservation of air volume and (6.8), integrate it from x to infinity where $j_a = 0$, and obtain

$$j_a(x) = -\partial_t \int_x^\infty \theta_w \, \mathrm{d}x' \ . \tag{6.11}$$

Inserting this back into (6.9) yields

$$j_w(x) = -K_w(\theta_w)\partial_x \psi_m - \frac{K_w(\theta_w)}{K_a(\theta_a)} \partial_t \int_x^\infty \theta_w \, \mathrm{d}x' \ . \tag{6.12}$$

We finally notice that (6.11) may also be inserted back into the second equation of (6.8) to define a modified matric potential that depends on $\partial_t \theta_w$. The additional term is sometimes referred to as the *dynamic matric potential* and is contrasted to the static matric potential (3.19) that corresponds to the pressure jump across the water-air interface.

6.1.3
Discontinuous Multiphase Regime

When the air content drops below a critical value, the continuity of the air phase is lost and a *residual air phase* forms. These air bubbles are completely enclosed by the water phase, either by bulk water or by water films that are adsorbed to the matrix. They typically occur in large cavities or pores. Air may leave such an isolated region through various processes, either as bubbles or by dissolution and subsequent diffusion. Bubbles form and leave the isolated region suddenly when the capillary pressure $p_w - p_a$ increases such that the interfacial radius becomes smaller than the largest radius of the blocking pore space, hence, when the air-entry pressure is exceeded. Dissolution occurs if the mole fraction κ_i of dissolved gaseous component i is less than the corresponding partial pressure $p_{a_i} = K_i \kappa_i$ in the isolated air phase, where K_i is the Henry constant for component i. This process thus leads to a characteristic shift of concentrations in the air- and in the water-phase.

As a consequence of bubble formation and dissolution, the air pressure in different isolated regions is generally different and formulations of the form (6.8) are quite useless. Instead, the pore-scale processes of bubble formation and dissolution must be averaged, a problem that has yet to be solved.

6.2
Stationary Flow

In the following, we only consider degenerate multiphase flow and refer to it as the *Richards regime*. The relevant material properties for rigid porous media are the soil water characteristic and the hydraulic conductivity that were already introduced and discussed in Section 3.4.

We first study the stationary flow of water in the vadose zone (i) as a limiting case for the more realistic transient flow and (ii) as an approximation at greater depths where higher frequency components are dissipated away. In a stationary system, all time derivatives vanish and the Richards equation (6.3) reduces to

$$-\mathsf{K}(\psi_m)[\nabla \psi_m - \rho_w \mathbf{g}] = \mathbf{j}_w \ , \tag{6.13}$$

i.e., to the statement that the Buckingham-Darcy flux (6.1) is constant in time. This problem is still hard to solve in general, since for heterogeneous media, $\mathbf{j}_w$ varies in space.

6.2.1
Uniform Soil

Consider a uniform and isotropic soil with a constant water table at depth $z = 0$ and vertical, constant water flux in the vadose zone. This leads to a one-dimensional problem which allows to greatly simplify (6.13).

We choose the z-axis to point downwards – hence $z < 0$ above the water table – use h_m instead of ψ_m, and obtain the ordinary differential equation

$$-K(h_m)[\mathrm{d}_z h_m - 1] = j_w^0, \tag{6.14}$$

where j_w^0 is constant. This may be transformed into

$$\mathrm{d}_z h_m = 1 - \frac{j_w^0}{K(h_m)} \ . \tag{6.15}$$

General Case The general form $h_m(z)$ for stationary flow may be deduced by recalling that $K(h_m)$ decreases monotonically, and rapidly, as h_m becomes more negative. Starting at the water table, at $z = 0$ and $h_m = 0$, we distinguish four regimes:

- *No flow* With $j_w^0 = 0$, the result is trivial, $h_m = z$, and independent of the hydraulic conductivity function.
- *Constant infiltration with* $0 < j_w^0 < K(0)$ The slope of $h_m(z)$ at $z = 0$ is positive but less than 1. Hence, h_m becomes more negative as z becomes more negative above the water table and $K(h_m)$ decreases. As $j_w^0/K(h_m)$ approaches 1, the slope decreases to 0 and h_m approaches a constant given by $K(h_m) = j_w^0$. In this limit of constant h_m, gravity is the only driving force of the flow which is consequently referred to as *gravity flow*.
- *Constant infiltration with* $j_w^0 \geq K(0)$ The slope at $z = 0$ is 0 or negative, hence h_m increases as z decreases, the medium remains saturated, and K remains constant. Again, (6.15) can be solved trivially and yields the linear function $h_m(z) = [1 - j_w^0/K(0)]z$, which is in analogy to groundwater flow.
- *Constant evaporation with* $j_w^0 < 0$ The slope of $h_m(z)$ at $z = 0$ is larger than 1, $K(h_m)$ thus decreases rapidly. As a consequence, the slope increases ever more and eventually leads to a constant value of z, the maximum height to which the flux j_w^0 can be maintained by evaporation.

Exemplary Case A separation of variables in (6.15) leads to

$$\int_0^{h_m} \frac{\mathrm{d}h'_m}{1 - j_w^0/K(h'_m)} = \int_0^z \mathrm{d}z' = z \, . \tag{6.16}$$

This integral cannot be evaluated analytically for the Mualem-van Genuchten parameterization, however. Hence, we choose the similar but simpler form

$$K(h_m) = \frac{K_0}{1 + [\beta h_m]^m} \, , \quad \beta < 0, \; m > 1 \tag{6.17}$$

which yields

$$\frac{h_m}{1-\xi} \, {}_2\mathrm{F}_1\Big(1, \frac{1}{m}, 1 + \frac{1}{m}, \frac{[\beta h_m]^m \xi}{1-\xi}\Big) = z \, , \tag{6.18}$$

where ${}_2\mathrm{F}_1$ is the hypergeometric function and $\xi := j_w^0/K_0$. This turns into more familiar functions for specific values of m, e.g., for $m = 2$, into (Figure 6.2)

$$\frac{1}{\sqrt{\xi[\xi-1]}} \arctan\Big(\frac{\sqrt{\xi}}{\sqrt{\xi-1}} \beta h_m\Big) = -\beta z \, . \tag{6.19}$$

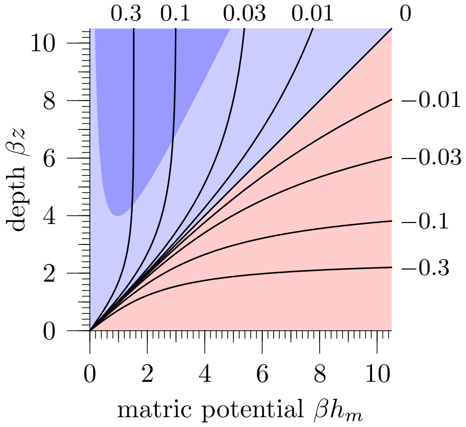


Figure 6.2. Dimensionless matric head βh_m as function of dimensionless height βz above constant water table for constant infiltration (blue) and constant evaporation (red) as given by (6.20) and (6.22), respectively. Each curve is labeled with the dimensionless flux $\xi = j_w^0/K_0$. The deeper shade of blue indicates the approximate domain of gravity flow obtained by choosing $\sqrt{\xi[1-\xi]}\beta z = 2$ for the argument of tanh in (6.20).

Infiltration With the water flux in the direction of the vertical axis, hence $\xi > 0$, and $\tan(\sqrt{-\xi}) = \mathrm{i}\tanh(\sqrt{\xi})$, (6.19) transforms into

$$\beta h_m(z) = \sqrt{\frac{1-\xi}{\xi}}\tanh\bigl(\sqrt{\xi[1-\xi]}\beta z\bigr)\ , \quad 0 \le \xi \le 1,\ \beta < 0\ . \tag{6.20}$$

For the regime of *gravity flow* – infiltration far above the water table, hence $\xi > 0$ and $z \to -\infty$ –, the matric head approaches the constant value

$$h_m^\infty(\xi) = \frac{1}{\beta}\sqrt{\frac{1-\xi}{\xi}}\ . \tag{6.21}$$

As we noticed above, the existence of gravity flow does not depend on the details of the parameterization of $K(h_m)$ but only on the fact that K decreases with more negative h_m.

Evaporation For evaporation, where $\xi < 0$, (6.19) transforms into

$$\beta h_m(z) = \sqrt{\frac{\xi-1}{\xi}}\tan\bigl(-\sqrt{\xi[\xi-1]}\beta z\bigr)\ , \quad -1 \le \xi < 0,\ \beta < 0\ . \tag{6.22}$$

As expected from the general discussion, a stationary evaporation flux from a constant water table can only be maintained to the maximal height

$$z^\infty(\xi) = \frac{\pi}{2\beta\sqrt{\xi[\xi-1]}} \tag{6.23}$$

which is proportional to β^{-1}, which in turn is proportional to the largest pore radius of the porous medium, and approximately proportional to the inverse of the dimensionless flux ξ. This result, which again hinges only on the sufficiently rapid decay of $K(h_m)$ not on its details, has important consequences for many natural phenomena.

Figure 6.3. Salt precipitation in arid mountain valleys near Qinghai Hu, China. Water from the infrequent rain events accumulates in the valley bottom from where it evaporates.

Example: Natural Vegetation Vegetation transpires water – evaporates it through leaves – in order to maintain the transport of nutrients from roots to leaves. For grasslands or forests, the corresponding flux is on the order of 1 $\mathrm{mm\,d^{-1}}$. In humid regions with frequent rainfall, this water is extracted from the rooted part of the vadose zone which in turn is replenished by precipitation. In more arid regions, rainfall is highly intermittent with high precipitation rates for very short times that are separated by extended dry periods. Such a regime favors groundwater recharge and vegetation has to recover the water from deeper layers. As (6.23) shows, this is possible through capillary rise, but only to some maximal height. This also explains the change in vegetation cover when groundwater levels change permanently, e.g., because of new extraction wells or because of water import from other region for irrigation.

Example: Salinization Upon evaporation, the salt load of the water is left near the soil surface and increases the concentration of the remaining soil solution. In humid regions, it is diluted and transported back to greater depths by frequent rainfall. In arid regions, however, concentrations may increase to levels that are not tolerable by vegetation and salt may actually precipitate to form a solid crust (Figure 6.3).

Salinization is the most challenging issue in irrigation agriculture, particularly in regions with a net water deficit. Irrigation water is either produced *in situ* from an underlying aquifer or it is imported from other catchments. In both cases, the salt carried by the water that is lost to evaporation remains in the soil's surface layer and accumulates there unless it is moved to deeper layers by regular flushing, which requires additional water. Optimal management aims at moving the salt beyond the root zone with the minimal amount of water.

A particularly severe problem occurs in irrigated regions with large water imports: As a result of the additional seepage, the water table rises and, with (6.23), the evaporation increases. This accelerates the salinization, sometimes dramatically. As a countermeasure, the water table must be lowered artificially with the water either being transported to some waste areas, like the Gobi and the Taklamakan deserts in China's Xinjiang province, or back into some river, as is done for instance in the North China Plain.

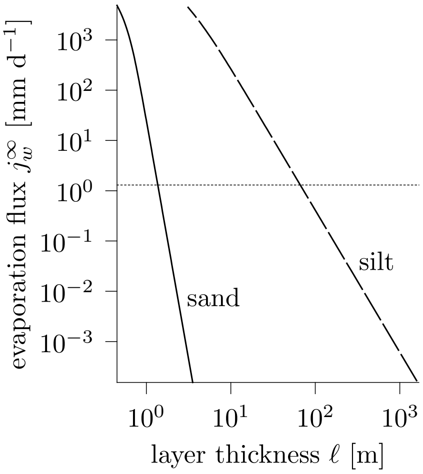


Figure 6.4. Maximum evaporation flux j_w^∞ that can be sustained through a layer of thickness ℓ for different materials. Calculated with (6.23) for parameters chosen such the conductivity function becomes asymptotically identical to that of the materials shown in Figure 6.6. The horizontal dotted line represents the mean global evapotranspiration rate is some 1.3 mm d^{-1}.

Example: Maximal Capillary Rise of Typical Soil Materials Instead of using (6.23), which is limited $m = 2$, we extend the integral in (6.16) to $-\infty$. This yields the maximum height $z^\infty(\xi)$ to which the dimensionless evaporative flux ξ can be maintained by a water table as

$$z^\infty(\xi) = -\frac{1}{\beta}\frac{\pi/m}{\sin(\pi/m)}\frac{[1+\xi]^{\frac{1}{m}-1}}{\xi^{\frac{1}{m}}} \ , \quad \beta < 0 \ . \tag{6.24}$$

Next, we choose the parameters of the conductivity function (6.17) such that it approaches the Mualem-van Genuchten function (3.56) asymptotically. Comparison with (3.57) shows that this is accomplished in the limit $\alpha h_m \gg 1$ with $\alpha = \beta$, $a[n-1] + 2n = m$, and with K_0 in (6.17) larger than that in (3.57) by the factor $[1-1/n]^{-2}$. Inserting the values from Table 3.1 leads to Figure 6.4. It shows that both soils are easily able to sustain very high evaporation fluxes, well beyond peak values of some 40 mm d^{-1} as they can occur in hot and dry climates, provided that the soil layer is shallower than about one meter. In such wet environments, the evaporation flux is not limited by the soil, but by the available energy for the evaporation of the water and by the relative humidity of the air. As the soil layers get thicker, the maximum evaporation rate drops very rapidly, particularly in the sand. Indeed, the mean global evapotranspiration rate of some 1.3 mm d^{-1} (Figure 1.1) can be sustained by the sand through of soil layer of just a little more than one meter. In contrast, the silt allows to maintain this flux through layers of up to some 80 m thickness. This corresponds with the experience that fine textured materials are more susceptible to salinization, while on the other hand a layer of very coarse material, so-called mulching, dramatically reduces evaporative losses in your home garden, and the coarse Saharan sand protects the groundwater that originated in the ice ages from evaporation.

6.2.2
Layered Soil

Consider a soil that is composed of uniform, isotropic, and horizontal layers and assume a stationary, uniform, and vertical flow. For stable flow, this again reduces to the one-dimensional problem (6.14).

Boundaries The water flux is continuous across the boundary between two layers. As a consequence, the water potential ψ_w is continuous and with it also the matric potential ψ_m and the matric head h_m. On the other hand, the soil water characteristic $\theta(\psi_m)$ is in general different for the two layers, hence θ is in general discontinuous.

Equating the water fluxes (6.14) on both sides of the interface leads to

$$\frac{K^-(\theta^-)}{K^+(\theta^+)} = \frac{\mathrm{d}_z h_m^+ - 1}{\mathrm{d}_z h_m^- - 1} \, , \tag{6.25}$$

where the superscripts − and + indicate the values above and below the interface, respectively. Since the conductivity will in general be different in the two layers, the matric head will in general have a kink, i.e., h_m is continuous, but its derivative is not.

Effective Conductivity For one-dimensional flow perpendicular to the layers, the effective conductivity is obtained from writing the Buckingham-Darcy law (6.14) as

$$-\frac{1}{j_w}\frac{\mathrm{d}h_w}{\mathrm{d}z} = \frac{1}{K(z)} \, , \tag{6.26}$$

with the hydraulic head $h_w = h_m - z$. Integrating from z_0 to z_1 and dividing by $z_1 - z_0$ yields

$$-\frac{1}{j_w}\frac{h_w(z_1) - h_w(z_0)}{z_1 - z_0} = \frac{1}{z_1 - z_0}\int_{z_0}^{z_1} \frac{\mathrm{d}z}{K(z)} =: \frac{1}{K^{\mathrm{eff}}} \tag{6.27}$$

or, for two layers with thicknesses ℓ_1 and ℓ_2 and constant conductivities K_1 and K_2, respectively,

$$\frac{\ell_1 + \ell_2}{K^{\mathrm{eff}}} = \frac{\ell_1}{K_1} + \frac{\ell_2}{K_2} \, . \tag{6.28}$$

These results are expected, of course: defining the specific hydraulic resistance $R := 1/K$ we recognize that the effective resistance equals the mean resistance along the flow path.

Two-Layer Soils A number of natural soils may be approximated as two-layer systems. Examples are sedimentary environments where large braided rivers created thick deposits of gravels and coarse sands during melting phases of past ice ages. With the greatly reduced flow of the subsequent warm phase, the erosional power of the rivers decreased and deposits consisted of finer grained material that ranges from fine sand all the way to clay. The

Figure 6.5.
Typical two-layer soil, here on the alluvial fan of Neckar river near Heidelberg, Germany. The top layer, down to some 1.5 m, consists of a loamy sand and the bottom layer of stones and gravel in a coarse sand matrix. The transition between the two main layers is very sharp, just a few centimeters.
Apparently, the main layers are not uniform themselves but consist of sub-layers. The top unit may be differentiated into (i) the plough layer down to about 0.2 m which is darker because of the higher organic matter content, (ii) a lighter colored horizon down to some 0.8 m which has a slightly higher sand content, (iii) a marble textured transition layer down to 1.2 m where iron and manganese were dissolved during times of high water saturation and ensuing low redox potentials in some places (whitish), transported, and precipitated at places with high redox potential (brownish), and finally (iv) a dark colored layer with a significantly increased clay content right above the sand-gravel unit. Also this last visible unit exhibits multiple layers of fine and coarse material.
The top layer was frozen at the time when the picture was taken, thence the different appearance. The differently colored units on the scales are 0.1 m long.

transition between corresponding layers is often rather sharp (Figure 6.5). Two-layer structures are typically also found in anthropogenically modified soils. Most prominent are plough layers where the pore structure is completely different from the underlying parent material due to frequent mixing and higher organic matter content from decaying roots.

The key to understanding stationary flow is the non-linear ordinary differential equation (6.14) which formulates that the Buckingham-Darcy flux is constant. Although the analytic approximations we employed for the uniform soil could be extended to two-layer soils, this gets clumsy. We thus choose to transit to numerical solutions which offer a much greater flexibility. Such a solution involves (i) integrating (6.15) numerically to obtain $h_m(z)$ and (ii) calculating $\theta(z)$ and $K(\theta)$ from the respective material properties. Since the integration is done numerically, we may also use more complicated representations of $K(h_m)$, for instance the Mualem-van Genuchten parameterization. Such flexibility comes at the price of a much less general result, however. While analytical solutions are valid for entire classes of parameter functions and yield classes of solutions, thereby facilitating deeper insight,

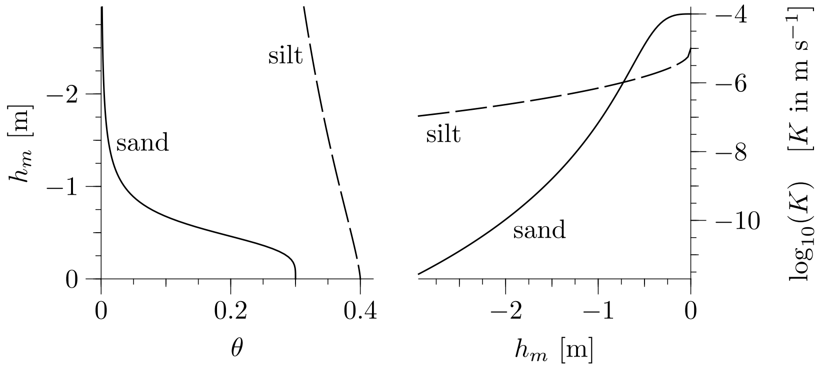


Figure 6.6. Mualem-van Genuchten parameterizations of the soil hydraulic properties used for Figures 6.7–6.12. Parameter values are given in Table 6.1. The conductivity functions intersect at $h_m^\times = -0.730$ m and $k^\times \approx 1.01 \cdot 10^{-6}$ m s^{-1}. Notice the linear scale used for h_m.

every change in a parameter's value necessitates a new numerical solution. Still, when it comes to specific results, we realize that also the very evaluation of an analytic solution requires often highly specialized numerical methods. More fundamental, analytic solutions become outright impossible already for moderately complicated problems.

In the following, we choose two materials, call them sand and silt, and construct the more complicated architectures from them. Let their hydraulic properties be described by the Mualem-van Genuchten model (3.44) and (3.56) with parameters given in Table 6.1 and illustrated in Figure 6.6. As we found earlier – Figure 3.26 on page 68 –, for small values of the matric head h_m, i.e., near saturation, the conductivity of the sand is much larger than that of the silt but then drops quickly as h_m becomes more negative. The two conductivity functions intersect at $h_m^\times = -0.730$ m. This corresponds to a water content θ of 0.083 in the sand and of 0.378 in the silt.

Consider two stacked horizontal layers of sand and silt, each 1 m thick. The effective conductivity of this formation at water saturation is obtained from (6.28) as $K_0^{\text{eff}} = 1.818 \cdot 10^{-5}$ m s^{-1}. This will be the reference value for the water flux since it is the maximum flux that can be sustained without

Table 6.1. Mualem-van Genuchten parameters for the soil hydraulic properties shown in Figure 6.6 and used for Figures 6.7–6.12. For both materials, $\theta_r = 0$ and $a = 1/2$ is chosen.

	θ_s	α [m^{-1}]	n	K_0 [m s^{-1}]
sand	0.3	−2.0	4.00	$10 \cdot 10^{-5}$
silt	0.4	−0.5	1.33	$1 \cdot 10^{-5}$

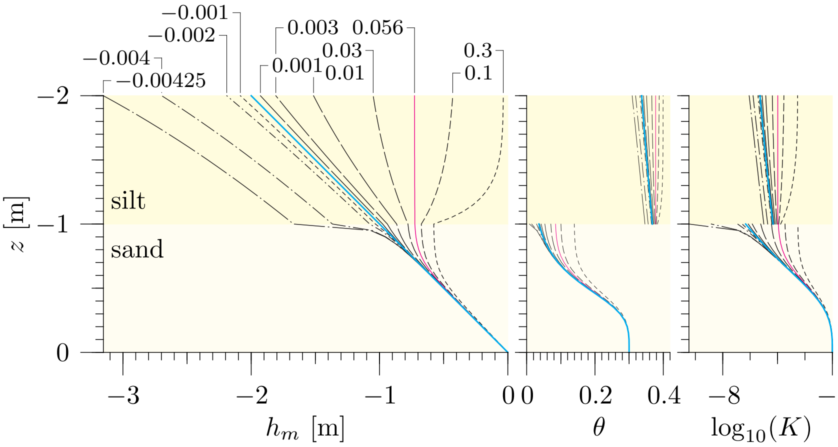


Figure 6.7. Matric head, water content, and hydraulic conductivity for stationary infiltration and evaporation in two-layer profile. Corresponding curves have the same dash pattern. Numbers in the leftmost graph indicate the relative flux j_w/K_0^{eff}, where $K_0^{\text{eff}} = 1.818 \cdot 10^{-5}$ m s^{-1} (65.5 mm h^{-1}). Notice (i) the curves of static equilibrium (cyan) which separate infiltration from evaporation regimes and (ii) the flux for which K is continuous across the layer interface (magenta) which separates the regimes where sand is a better/worse conductor than silt. Also notice the differences in the axis between the leftmost plot and Figure 6.2.

ponding. Finally, we choose the origin of the downward pointing z-axis at the constant water table at the lower end, 2 m below the soil surface.

Static Equilibrium With $j_w = 0$ in (6.15) $\mathrm{d}_z h_m = 1$, hence $h_m(z) = z$ independent of the material properties. The water content profile $\theta(z)$ is then equal to the soil water characteristic $\theta(h_m)$ as is illustrated by the cyan curve in Figure 6.7. For this case, the hydraulic variables may be calculated directly from the parameterizations of $\theta(h_m)$ and $K(h_m)$ or estimated from Figure 6.6. Values for some interesting depths are given in Table 6.2. We notice that for the chosen parameters and layer dimensions, the matric head at the interface is lower than $h_m^\times$, the intersection of the conductivity functions. Hence, for static equilibrium the conductivity of the sand near the interface is lower than that of the silt.

Table 6.2. Values for some hydraulic variables in two-layer soil, sand and silt, at static equilibrium for a water table at $z = 0$, hence with $h_m = z$. Each of the layers is 1 m thick.

	silt		sand	
	$z = -2$ m	$z = -1$ m	$z = -2$ m	$z = -1$ m
θ	0.337	0.368	0.005	0.036
K [m s^{-1}]	$2.29 \cdot 10^{-7}$	$6.88 \cdot 10^{-7}$	$1.06 \cdot 10^{-10}$	$6.83 \cdot 10^{-8}$

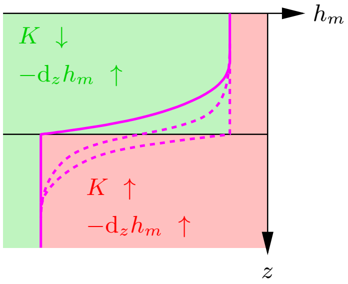


Figure 6.8.
Sketch for shape of $h_m(z)$ near an interface. In the green region, the two components of j_w can compensate each other such that $j_w = \text{const}$ during the transition between the two gravity flow regimes. In the red region, K and $-\mathrm{d}_z h_m$ both change in the same direction, j_w thus cannot remain constant as required. The transitions indicated by the dashed lines are thus not permissible.

The further discussion must distinguish between the two possible configurations, silt on top of sand and *vice versa*.

Silt Overlying Sand For infiltration $j_w > 0$ and, with (6.14), $\mathrm{d}_z h_m < 1$. As a consequence, throughout the profile h_m is higher than in static equilibrium and with it also θ and K (Figure 6.7). For the flux $j_w/K_0^{\text{eff}} \approx 0.056$, h_m at the interface equals $h_m^{\times}$. Then, by definition, the conductivities on both sides of the interface become equal and, with (6.25), $\mathrm{d}_z h_m$ is continuous. Notice though that θ remains discontinuous at the interface.

For higher infiltration fluxes, sand turns into a better conductor than the overlying silt. This is reflected in $h_m(z)$ which in the silt decreases towards the interface. If on the other hand the flux is lower, such that silt is the better conductor, then the matric head increases towards the interface. It is instructive to study these shapes in more detail. First consider the curve for $j_w/K_0^{\text{eff}} = 0.3$ in Figure 6.7. At sufficient distances from the interface, $h_m(z)$ becomes constant in both layers as expected for the regime of gravity flow. The transition from the constant matric head in the silt to the lower value in the sand creates a region where $\mathrm{d}_z h_m < 0$. Why does this transition occur entirely within the upper layer? With $\mathrm{d}_z h_m < 0$ the driving force of the water flux is stronger than in the gravity flow regime. Since the flux is constant – we consider stationary flow – the conductivity must be smaller in the transition region than in the gravity flow region. Hence also h_m must be lower. These two conditions can only be satisfied in the upper layer (Figure 6.8). Similarly for a lower flux, for instance $j_w/K_0^{\text{eff}} = 0.03$. The silt is now a better conductor than sand, hence $\mathrm{d}_z h_m > 0$ in the transition region. The driving force in this zone is thus smaller than in the gravity flow regime, hence K must be larger. Again, this can only be satisfied in the upper layer.

For evaporation $j_w < 0$, hence $\mathrm{d}_z h_m > 1$ and h_m is lower than in static equilibrium in the entire profile. Consequently, θ and K are everywhere smaller than for $j_w = 0$. As already found for the uniform medium in Figure 6.2, $\mathrm{d}_z h_m$ has to increase very rapidly with increasing evaporative flux in order to compensate for the rapidly decreasing conductivity. This effect is strongest in the sand layer which has a much lower conductivity near the interface than the silt. Indeed, the driest point in this two-layer soil is always the upper end of the sand. Still, this soil can sustain a maximum

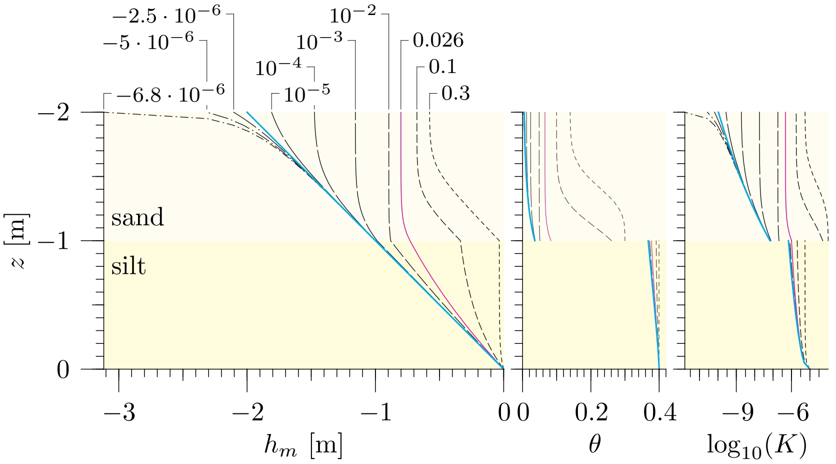


Figure 6.9. Same as Figure 6.7 but for interchanged layers, sand on top of silt.

evaporative flux $j_w/K_0^{\text{eff}} \approx -0.004$ which corresponds to some 0.26 mm h^{-1}. It can thus easily cover the water demand of any vegetation, even if it is shallow-rooted. On the other hand, in an arid climate it would become saline in a short time due to the high evaporative loss. We notice in passing that such an evaporative flux consumes some 0.18 kW m^{-2} which reduces the sensible heat flux and thus leads to a strong cooling of the surface. Obviously, all these values are highly dependent on the thickness of the sand layer. Just changing the water table by one meter, which happens easily under irrigated land, would change the maximum evaporative flux by 1...2 orders of magnitude, depending the direction of the change, with obvious consequences for vegetation and salinization.

Sand Overlying Silt The phenomenology of this layering is quite different from the previous one: The top layer is generally very dry, even for rather high infiltration fluxes, whereas the bottom layer is always nearly saturated (Figure 6.9).

For an infiltration flux that exceeds 1.7 mm h^{-1} ($j_w/K_0^{\text{eff}} = 0.026$), the silt layer becomes limiting to the flow while for lower fluxes, the sand is limiting. This is nicely indicated by the shape of $h_m(z)$ which is characteristically different from that in Figure 6.7. The separation between the two regimes is at a smaller flux than for the reverse layering because the transition to the regime of gravity flow is wider for the silt than for the sand. Again, $h_m(z)$ can be understood qualitatively by considering the two components of the flux, $K(h_m)$ and $\mathrm{d}_z h_m - 1$, and their transition between the gravity flow regimes in the two layers.

For evaporation, the sand layer is the major obstacle. Indeed, h_m in the silt layer does not change significantly between no flux and maximum evaporative flux because the conductivity in the silt is at least an order of magnitude

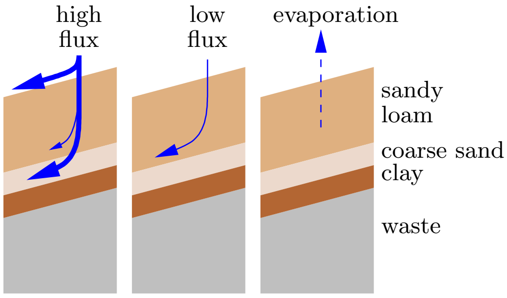


Figure 6.10. Schematic action of an inclined protective capillary barrier for a waste disposal site, for high and medium rainfall intensities and for evaporative regime.

higher than in the sand. Hence, $h_m(z)$ in the sand has the same shape as found in Figure 6.2 for a uniform medium above a water table. Further notice that the maximum evaporative flux in this configuration is much smaller than for the previous one, just $j_w/K_0^{\text{eff}} = -6.8 \cdot 10^{-6}$ corresponding to a maximum loss of some 3.9 mm y^{-1}. Even in a rather arid region, this loss can be compensated by rainfall, the groundwater would thus be conserved. Compare this with the loss of some 2.3 m y^{-1} that results from the reverse layering to appreciate the importance of the nonlinear dynamics. We recall in passing that in groundwater with its essentially linear dynamics the order of layers does not influence the overall flow.

Example: Capillary Barrier Two-layer structures are not only common in nature but they are also implemented in many engineering structures. An example of this is an inclined coarse-grained layer that is overlaid by a fine-grained layer. Such a setting may for instance be used to protect a waste disposal site from percolating water is sketched in Figure 6.10. Here a sandy loam overlays a coarse sand with the two forming a so-called *capillary barrier*. It is followed by a clay layer which blocks water due to its inherently low conductivity, even when fully saturated. Typically, the top layer carries some grass vegetation to counteract surface erosion and to further evaporation.

For high rainfall intensities, the conductivity of the top layer becomes limiting and the surplus generates surface runoff. The infiltrating water leads to a practically complete saturation of the sandy loam such that the matric head at the interface to the coarse sand increases such that the coarse layer becomes conductive. Due to its very high conductivity, usually a few orders of magnitude higher than that of the top layer, it can easily maintain the downhill flux of the percolating water. The clay layer prevents the breakthrough into the waste deposit and enforces the downhill flow.

For medium and low fluxes, the sandy loam is a much better conductor than the coarse sand and the water flows downhill above the interface. Of course, the water accumulates downhill and, if the barrier is too long, will eventually saturate the top layer and penetrate into the coarse sand. Here, it is again carried away very efficiently, however.

A further function of the coarse sand is as barrier against capillary rise during evaporation periods. Since it can sustain only exceedingly small upward fluxes, the underlying clay remains wet which prevents desiccation cracks. All the evaporating water stems from the top layer and increases its capacity for the next rainfall event.

6.2.3
Heterogeneous Soil

As a final approximation to natural architectures, consider a soil that consists of irregularly shaped layers (Figure 6.11). Let each of the layers be uniform and consist of one of the materials sand or silt studied above. We restrict our study to a two-dimensional structure since three-dimensional simulations are very expensive due to "the curse of dimensions", their results are difficult to represent, and they lead to the qualitatively same results for the essentially single-phase flow we consider here. We furthermore assume periodic vertical boundaries, i.e., water that leaves through one boundary enters through the other. To accommodate these boundaries, also the heterogeneity has been generated with such a periodicity. Such periodic boundary conditions are important in heterogenous architectures to prevent artifacts from impermeable boundaries. Finally, we again impose a constant water table at $z = 0$ and choose the soil surface at $z = -2$ m.

Notice in passing that along with an increasing complexity of the architecture goes the requirement for more powerful tools. Stationary flow can be studied analytically in uniform media while one-dimensional layered media already require numerical methods. These may be rather simple ones like Runge-Kutta integration for coding yourself [*Press et al.* 2002] or generic tools like *Mathematica* and *GNU Octave* for quickly obtaining accurate solutions. In contrast, heterogeneous media require a full-blown numerical Richards solver.

Infiltration We envisage a long-lasting constant rain and prescribe a uniform and constant infiltration flux through the upper boundary. The matric head at the surface will then in general not be constant but will adjust itself such that the impressed flux is maintained.

Low Infiltration Flux Consider the flux $j_w^0 = 1.16 \cdot 10^{-8}\ \mathrm{m\,s^{-1}}$ ($1\ \mathrm{mm\,d^{-1}}$) which corresponds to a very weak precipitation as it is typical for fall-out from fog. For a gravity flow regime in a uniform sand, this would lead to the matric head $h_m = -1.21$ m and the saturation $\Theta = 0.068$. In the silt on the other hand the corresponding values would be $h_m = -7.56$ m and $\Theta = 0.62$. Notice however, that with the soil surface 2 m above the water table and with infiltration, $h_m \geq -2$ m at the surface. Hence, the silt will be far from the regime of gravity flow and can easily sustain the imposed infiltration flux without the need to increase Θ appreciably. In contrast, for the sand gravity flow is a reasonable approximation near the surface and, in order to sustain the flux, Θ has to increase by more than a factor of 4 from its static equilibrium value of 0.016.

Inspection of the data shows that the average matric head in the top 0.5 m of the heterogeneous soil shown is $\langle h_m \rangle = -1.55$ m (Figure 6.11). For this value, the conductivity of the silt is more than two orders of magnitude higher than that of the sand. This is also manifest in the streamlines which

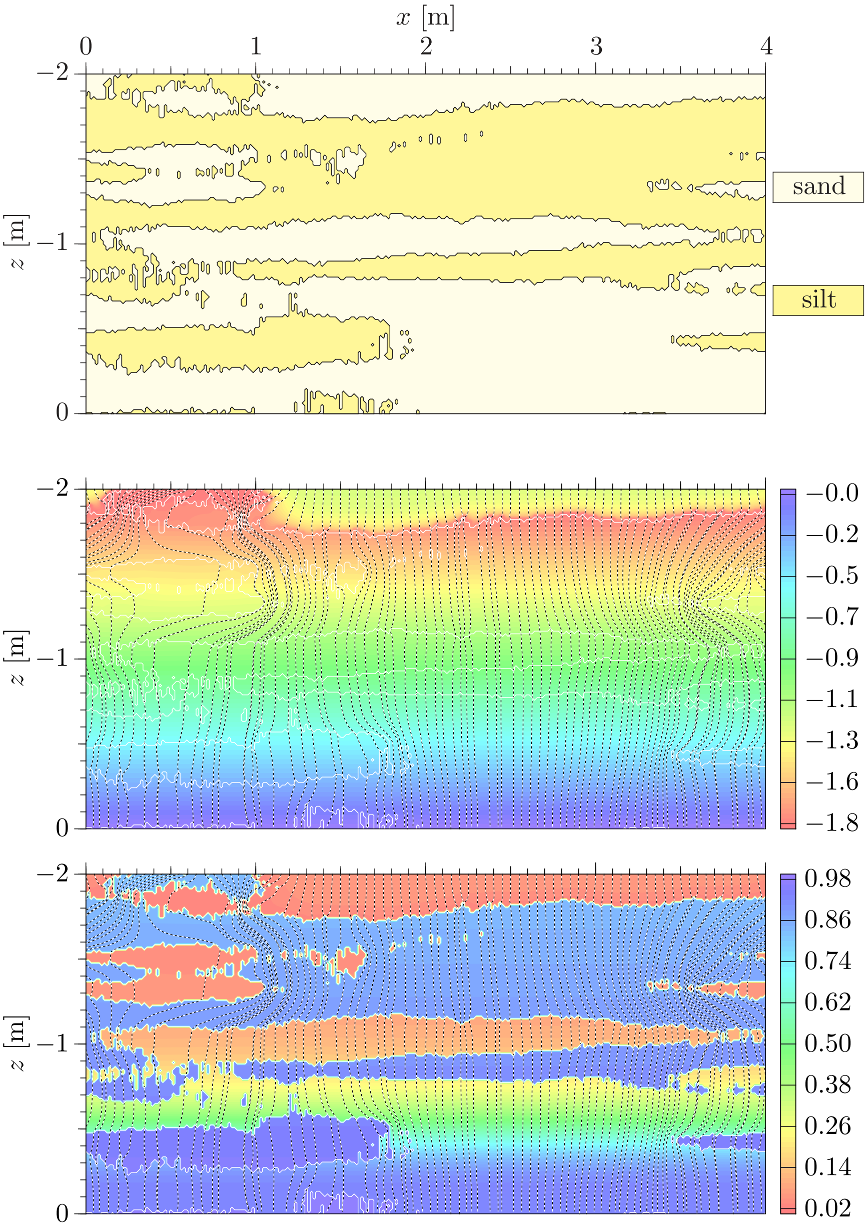


Figure 6.11. Matric head h_m [m] (**middle**) and water saturation $\Theta = \theta/\theta_s$ (**bottom**) for a heterogeneous soil composed of irregular sand and silt layers (**top**) for the infiltration flux $j_w^0 = 1.16 \cdot 10^{-8}$ m s^{-1} (1 mm d^{-1}). White lines in the lower two graphs represent the boundary between the materials, dashed lines are equidistant streamlines. Vertical boundaries are periodic.

tend to eschew the sand layers in the top part of the profile. With the matric head increasing towards the water table, the conductivity of the sand increases more rapidly than that of the silt. At $z = 1.27$ m, the mean matric head equals -0.73 m, the value where the conductivities of the two materials match. For greater depths, sand is a better conductor. This is again corroborated by the streamlines.

We further notice from Figure 6.11 that the matric head is rather constant in horizontal planes. This is akin to the predominantly vertical isolines ofh_m in groundwater flow, e.g., Figure 5.9 on page 112. An exception to this is the horizontally extended sand layer at the surface where the imposed uniform flux at the boundary, in conjunction with the greatly differing conductivities of the materials, leads to strongly nonuniform hydraulic gradients. At greater depths, these are attenuated because the flux-field adjusts and becomes nonuniform. As mentioned earlier, the smoothness is a consequence of Buckingham-Darcy's law which enforces h_m to be continuous.

In contrast, the water saturation Θ is discontinuous across all interfaces between different materials, as already found for the one-dimensional layered medium. This qualitative difference has some important implications: (i) Measurements of the water content reveal much about the arrangement of different materials. This is for instance exploited for the exploration of subsurface structures with ground-penetrating radar. This method maps the dielectric structure of the subsurface which is dominated by the water content. (ii) The water content at any particular time yields only very limited information on the dynamics of the water phase and on the current state of the site. To extract this, the matric potential would be a more useful variable which, alas, is much more difficult to monitor.

High Infiltration Flux Consider the infiltration flux $j_w^0 = 5.55 \cdot 10^{-6}$ m s^{-1} (20 mm h^{-1}) which corresponds to a heavy rain shower (Figure 6.12). For a uniform sand in the regime of gravity flow, this would lead to $h_m = -0.577$ m and $\Theta = 0.466$. The corresponding values for silt are $h_m = -0.032$ m and $\Theta = 0.999$. Hence, a uniform silt would be all but saturated to maintain the water flux which is about half of the its saturated conductivity. In contrast, the conductivity of the sand for $h_m = -0.032$ m would be by a factor of 30 larger than that of the silt. In the heterogeneous medium, the flow adjusts to this as is again reflected in the streamlines which tend to bypass the silt. Inspection of the data in the top 0.5 m of the heterogeneous medium yields an average value for the matric head in the sand of -0.22 m, considerably higher than expected for a uniform medium. Correspondingly, the average matric head in the silt is lower, -0.17 m instead of the -0.032 m expected for a uniform medium.

Finally, we notice that the silt is almost uniformly saturated throughout the entire profile. Inspection of the data actually yields Θ in the range 0.934...0.976. In contrast, characteristic saturation profiles develop in the sand layers which are determined on the one hand by the non-uniform flow

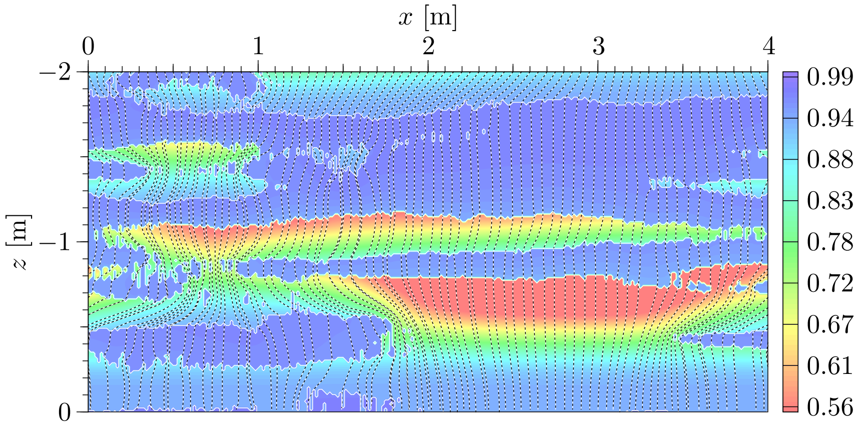


Figure 6.12. Water saturation Θ in the soil of Figure 6.11 for the high infiltration flux $j_w^0 = 5.56 \cdot 10^{-6}$ m s^{-1} (20 mm h^{-1}). Notice the different color scale.

from the region above and on the other by the ponding back from the less permeable silt layer below.

Evaporation A prime physical issue with evaporation is the appropriate boundary condition. Focussing on evaporation from some soil surface, we notice that two conditions must be met: (i) the energy for evaporating the water, some 2.5 MJ kg^{-1}, must be provided and (ii) the water must actually be removed by the atmosphere. The energy comes to a large part from solar radiation and to a typically smaller part through heat flow from the ground. Removal of water is by diffusion through a thin soil-atmosphere layer which is just a few millimeters thick and then through mostly turbulent transport in the atmosphere. Apparently, the situation is quite complicated with two compartments, soil and atmosphere, and two processes, water and heat, closely linked.

For a heterogeneous medium and with a focus on water flow within the soil, a reasonable approximation to the physical boundary condition is a uniform matric head at the soil surface which represents the dry atmosphere that drives evaporation. We prescribe in the following a uniform and constant matric head at the upper boundary, $h_m^0 = -6$ m (Figure 6.13). This value corresponds to a relative humidity of 99.95%, as may be calculated with (8.15) on page 247, hence to a very wet atmosphere. When interpreting this value, we must be aware that in reality, this would not be the humidity in the atmosphere, say a few meters above the soil surface. It corresponds to the humidity at the lower end of the very thin soil-atmosphere interface. More typical values for a moderately dry atmosphere is a relative humidity of 50% which corresponds to a matric head of some $-9.6 \cdot 10^3$ m. This discrepancy reflects the slow diffusive transport of water vapor through the interface.

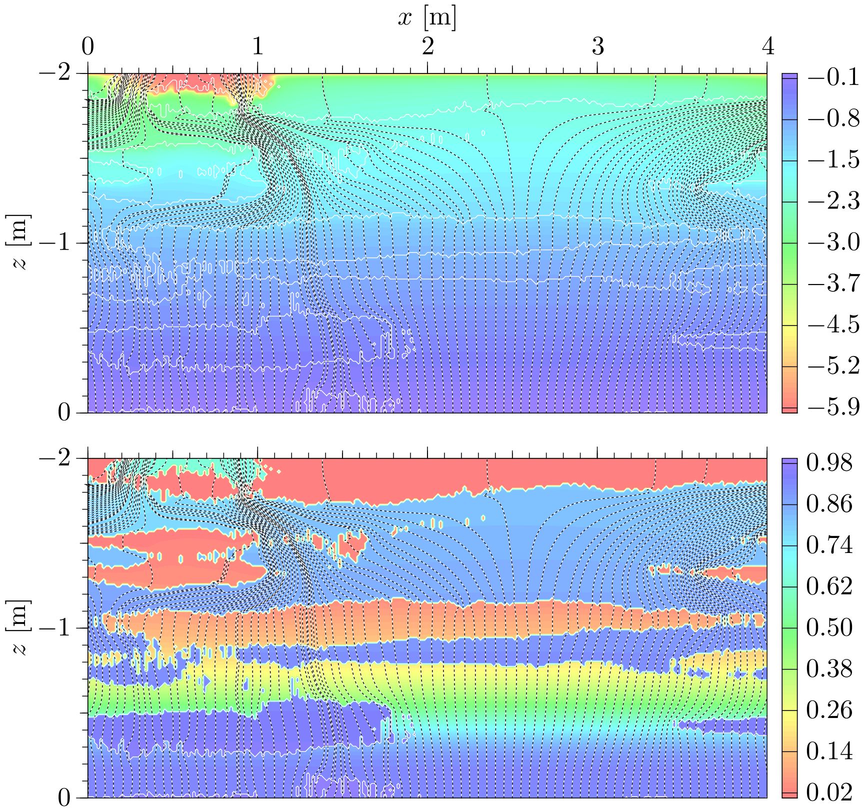


Figure 6.13. Matric head h_m [m] (**top**) and water saturation $\Theta = \theta/\theta_s$ (**bottom**) for stationary evaporation from the soil shown in Figure 6.11. At the upper boundary, $h_m = -6$ m which leads to the average flux $\langle j_w^0 \rangle = -3.55 \cdot 10^{-8}$ $\mathrm{m\,s^{-1}}$ (-3.06 $\mathrm{mm\,d^{-1}}$).

As already found in the one-dimensional study, the sand layer at the surface is a very effective evaporation barrier. Indeed, its conductivity for $h_m = -6$ m is some seven orders of magnitude lower than that of the silt. This leads to a very steep gradient right at the surface with the prescribed value of $h_m = -6$ m barely visible in the upper frame of Figure 6.13. This steep gradient does not penetrate far into the sand layer, however, such that only a very small flux can be maintained. This is corroborated by the few streamlines that traverse this layer.

Further focussing on the streamlines, we find that water rises uniformly out of the water table, as expected. A strong redistribution then occurs in the silt layer between 1.1 and 1.7 m above the water table. This indicates that the sand layer immediately below, between 0.9 and 1.1 m above the water table, is not a hydraulic obstacle, yet. However, sand higher than about 1.5 m above the water table is next to impenetrable and almost all of the

Figure 6.14.
The pattern of precipitated salt highlights the surface area through which most of the evaporation flux is lost. Image near Qitedaban, 4960 m asl, Xinjiang Autonomous Region, China.

rising water passes through two narrow ports some 0.2 m below the surface. This general behavior – uniform rise to intermediate heights, redistribution well below the soil surface, and eventual evaporation from the fine-textured materials – is the characteristic of evaporation from heterogeneous media. The details, of course, depend strongly on the specific architecture, including its hydraulic properties.

Looking more quantitatively, inspection of the data yields an average evaporative flux of $3.55 \cdot 10^{-8}\ \mathrm{m\,s^{-1}}$ ($3.06\ \mathrm{mm\,d^{-1}}$), a value that may in nature be expected for a warm spring day. The range of the evaporative flux across the surface is $[0.25 \ldots 7.9] \cdot 10^{-8}\ \mathrm{m\,s^{-1}}$. Such highly localized regions where evaporation from the surface occurs often become manifest through the precipitated salt and may then be observed in the field Figure 6.14.

6.2.4
Effective Dynamics

We may ask if all the small-scale details of the vadose zone need to be represented in order to obtain a faithful model for the flow of water or if an appropriately chosen "coarse-grained" description would also reproduce the pertinent phenomena. We already touched upon this issue for groundwater flow in Section 5.4.3. There, it sufficed to focus on the value of the hydraulic conductivity since there was no reason to question the representation of the essentially linear dynamics described by (5.1). The situation in the vadose zone is quite different since it is not obvious that simple averaging leads to useful representations of the highly non-linear dynamics.

As a first step, we need to illuminate the applicability of Richards equation (6.2). We recall that it rest upon three pillars, (i) the existence of an REV for the state variables θ and h_m, which implies the equilibrium between microscopic and corresponding macroscopic properties at the scale of the REV, (ii) the conservation of mass, and (iii) an empirical flux law where the macroscopic water flux is an instantaneous function of the gradient of the gradient of the macroscopic water potential. While the full problem has not been solved yet, not even for stationary flow, some results are available for special cases. In particular, *Neuweiler and Eichel* [2006] demonstrated that for vertical water flux in a soil with periodic thin horizontal layers, an effective representation in the general form of Richards' equation remains valid. For the case of dominating capillary forces, where $|\nabla h_m| \gg 1$, they furthermore found that the effective soil water characteristic is the arithmetic mean of its microscopic analogon and that the effective conductivity function is the corresponding harmonic mean, hence

$$\theta(h_m)_{\text{eff}} = \left\langle \theta^\mu(h_m^\mu) \right\rangle \quad \text{and} \quad K(h_m)_{\text{eff}} = \left\langle K^\mu(h_m^\mu)^{-1} \right\rangle^{-1} , \tag{6.29}$$

where the superscript μ indicates a microscopic quantity and $h_m^\mu = h_m$ to a first order approximation. The latter is expected as a consequence of the local equilibrium requirement and of the fact that $-\nabla h_m$ is a local driving force for the water flux. For the case where capillary forces are not dominating, *Neuweiler and Eichel* [2006] found that a generalization of Richards equation still represents the macroscopic dynamics. The required generalization allows different conductivity functions for capillarity and gravity as drivers of the water flux. While this hints at the eventual deterioration of Richards' equation, *Neuweiler and Eichel* [2006] found that (6.29) remains a useful approximation even if capillarity is not dominating gravity greatly.

The general situation in soils is considerably more difficult than what was covered above. Most importantly, soils typically consist of only a few dominating layers between the soil surface and the groundwater. In addition, gradients are often very high, particularly near the surface through the strong atmospheric forcing, and they change on rather short time scales which accentuate the local equilibrium requirement. In the following, we look into some of these issues for the still idealized situation of stationary flow but for the "nature-inspired" architecture studied in the previous section. To this end we interpret the simulation results in the context of an effective one-dimensional representation that is obtained by horizontally averaging the corresponding two-dimensional quantities.

Infiltration The averaged matric head $\langle h_m \rangle$ has a very similar shape as in a uniform medium (Figure 6.15). There is a regime of approximate gravity flow at sufficient distance from the water table and $\mathrm{d}_z \langle h_m \rangle \approx 1$ near the water table, as indicated by the dashed line. The only significant deviation from this simple shape occurs for the low infiltration flux near the surface. It results from the adjustment of the uniform infiltration flow to the heterogeneous hydraulic structure.

A sharp contrast to the regular behavior of $\langle h_m \rangle$ is the wildly fluctuating average volumetric water content $\langle \theta \rangle$. This is a manifestation of the predominantly horizontal layering and of the discontinuous nature of θ. We can thus easily identify the zones of dominant sand layers: the dry depth intervals. The fluctuation of $\langle \theta \rangle$ already hints at the difficulty to use it as a variable in a coarse-grained description, in an effective soil water characteristic or as independent variable for the conductivity function.

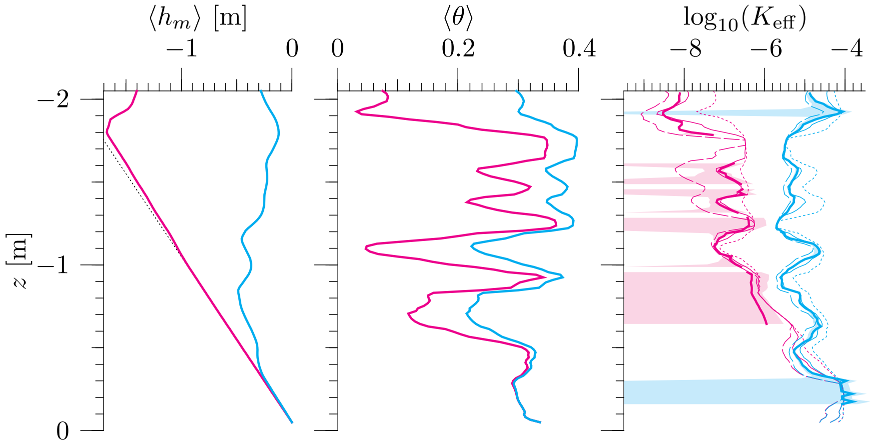


Figure 6.15. Average and effective quantities for a one-dimensional representation of the heterogeneous layers of Figure 6.11 for low (**magenta**) and high (**cyan**) infiltration flux. Matric head $\langle h_m \rangle$ and volumetric water content $\langle \theta \rangle$ are shown as horizontally averaged quantities (arithmetic mean). Conductivities are calculated from the two-dimensional fields with (6.30) (thin dotted lines), with (6.31) (thin dashed lines), and with (6.32) (thin solid lines), where averaging is again in the horizontal. Effective conductivities obtained from (6.33) are shown as thick lines with bands one standard deviation wide indicating the uncertainty. Sections of the curves with a very large uncertainty are not shown.

Of prime interest is the effective hydraulic conductivity and how to estimate it from the detailed two-dimensional information. We consider three candidate models, (i) continuous vertical stream-tubes,

$$K_{\text{eff}} = \langle K \rangle \ , \tag{6.30}$$

(ii) a perfectly layered medium with flow perpendicular to the layers,

$$K_{\text{eff}} = \left[\langle K^{-1} \rangle\right]^{-1} \ , \tag{6.31}$$

and (iii) the macroscopically uniform heterogeneous medium studied in Section 5.4 for which (5.4) yields the geometric mean

$$K_{\text{eff}} = \exp\bigl(\langle \log(K) \rangle\bigr) \ . \tag{6.32}$$

In all these expressions, K refers to the conductivity in the “fine-grained” (two-dimensional) representation and $\langle \cdot \rangle$ indicates arithmetic averaging in horizontal direction. Clearly, the assumptions of neither of these models are satisfied so they will offer rough estimates at best.

In the next step, we define the “true” effective hydraulic conductivity by postulating the one-dimensional form of the Buckingham-Darcy law (6.1) as $j_w^0 = -K_{\text{eff}}[\mathrm{d}_z \langle h_m \rangle - 1]$, where j_w^0 is the infiltration flux imposed at the upper boundary. Assuming the expression in brackets does not vanish, this leads to

$$K_{\text{eff}} = -\frac{j_w^0}{\mathrm{d}_z\langle h_m\rangle - 1} \; . \tag{6.33}$$

This expression has the advantage that it provides an implicitly flux-weighted estimate. Obviously, it becomes increasingly uncertain as $\mathrm{d}_z\langle h_m\rangle$ approaches 1 which is the case as the flux decreases and as the water table is approached. We notice that such a definition is only reasonable because $\langle h_m\rangle$ is a rather smooth function. Still, the objectivity of this definition would have to be ascertained. On the other hand, if an effective description exists at all, then (6.33) may be expected to yield the correct value for the effective parameter. Finally, it should be realized that all three approaches require the solution of the fine-grained problem. They are thus only useful to study the relation between fine-grained and coarse-grained formulations and to assess the feasibility of the latter.

For the high-flow regime with a high water saturation, (6.31)–(6.33) all yield reasonably similar estimates for $K_{\text{eff}}(z)$ even though the values differ by up to a factor of 4. The model of parallel tubes is clearly inferior and generally leads to a significant overestimation.

For the low-flow regime with its widely varying water content, the estimates differ quite significantly, often by more than an order of magnitude. They converge, as expected, at depths where one material forms a continuous layer since then the range of conductivities is much smaller.

Evaporation For infiltration, the heterogeneous model (6.32) and the one with horizontal layers, (6.31), were found to bracket the true conductivity, with (6.32) generally leading to a rather strong underestimation. In contrast, the stream-tubes model was found to be quite far off. These findings should not be generalized, however. They hinge on the specific structure of the fine-grained representation as well as on the specific forcing. This is demonstrated by the evaporation regime (Figure 6.16).

With evaporation, the matric head decreases very rapidly towards the surface. With this, the ratio between the hydraulic conductivity in the silt and the one in the sand increases very strongly. Consequently, the models (6.30)–(6.32), which weigh positive and negative deviations from the mean differently, yield estimates that differ by several orders of magnitude. Furthermore, the best estimates in the top 0.2 m are produced by the stream-tubes model, which fared worst in the infiltration regime. Looking at the streamlines in Figure 6.13 reveals the reason for this. The flow is strongly focused into the silt and the arithmetic averaging gives very little weight to the exceedingly low conductivity values of the sand. Overall however, for depths beyond the top 0.2 m, the heterogeneous model (6.32) again provides the most reliable values among the direct estimators.

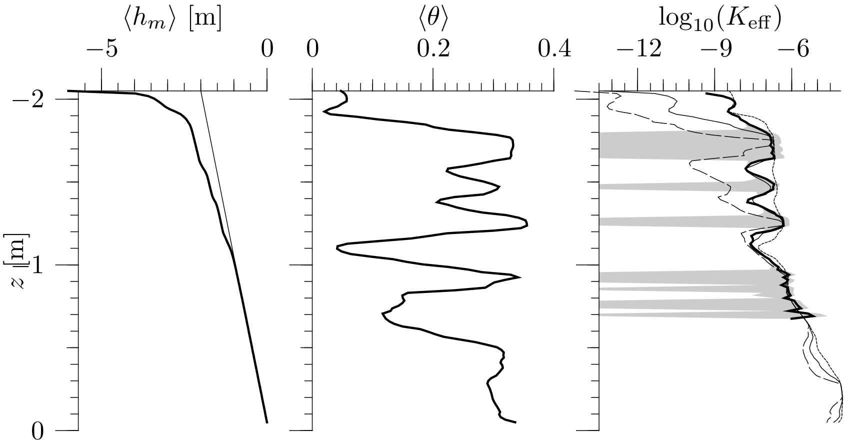


Figure 6.16. Same as Figure 6.15 but for evaporative flow shown in Figure 6.13. Notice the different scales.

6.3
Transient Flow

Water flow is predominantly driven by rainfall, evaporation from the soil surface, and root water uptake by plants. This is often referred to as *atmospheric forcing*. In passing, we recall that the soil-atmosphere system is strongly coupled and the atmosphere is driven by the soil as well, both through the evaporating water and the raising sensible heat (Figures 1.1–1.2). Accordingly, in a more comprehensive perspective, the two systems are considered together. Here, we are only interested in vadose zone flow and thus assume either the water flux or the matric potential to be given at the upper boundary as a function of time. A further though often less prominent driver of water flow in the vadose zone is a fluctuating water table of an underlying phreatic aquifer.

Atmospheric forcing leads to a highly irregular dynamics through its succession of rainfall events that are separated by evapotranspiration periods. Obviously, the characteristics of this forcing, in particular the average net flux and the degree of intermittency, depends on the climate region. As a first step to understanding such regimes, we study individual infiltration and drainage events which will be found to exhibit distinctly different features. This may be understood qualitatively through the Richards equation formulated in the θ-form (6.4)–(6.5) for a uniform porous medium. For the one-dimensional case, it may be written as

$$\partial_t \theta + \underbrace{V_w(\theta)\partial_z \theta}_{\text{advection}} - \underbrace{\partial_z \left[D_w(\theta)\partial_z \theta \right]}_{\text{dispersion}} = 0 \tag{6.34}$$

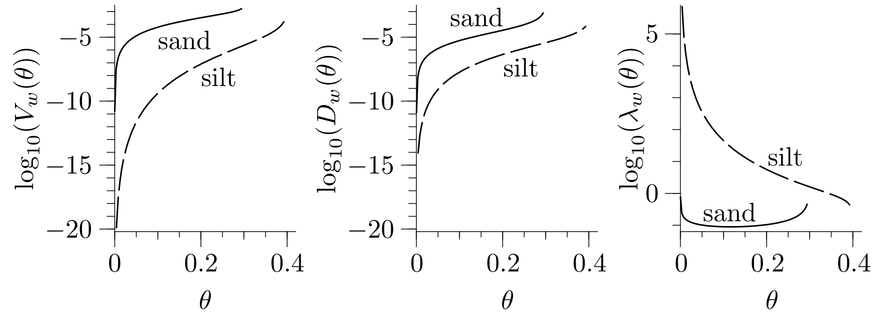


Figure 6.17. Soil water velocity $V_w(\theta)$ [$\mathrm{m\,s^{-1}}$] and diffusivity $D_w(\theta)$ [$\mathrm{m^2s^{-1}}$], defined by (6.5), together with soil water dispersivity $\lambda_w(\theta) := D_w(\theta)/V_w(\theta)$ [m], for the Mualem-van Genuchten parameterization of the two materials sand and silt considered above. Parameter values are given in Table 6.1.

with

$$V_w(\theta) := \frac{\partial K_w}{\partial \theta} \;, \qquad D_w(\theta) := \frac{\partial h_m}{\partial \theta} K_w(\theta) \;, \tag{6.35}$$

where K_w is expressed in units of $\mathrm{m\,s^{-1}}$. As mentioned before, this equation describes the movement of water as a superposition of advection and dispersion. A macroscopic volume of water with water content θ hence moves with velocity $V_w(\theta)$. This may be seen by dropping the dispersion term for the time being and considering the trajectory of that volume, the path $z_\theta(t)$ on which θ is constant, hence $\mathrm{d}\theta = 0$. Inserting $\theta(z(t), t)$ leads to

$$\mathrm{d}\theta = \left[\partial_z \theta \frac{\mathrm{d}z}{\mathrm{d}t} + \partial_t \theta\right] \mathrm{d}t \;, \tag{6.36}$$

which vanishes only if $\frac{\mathrm{d}z}{\mathrm{d}t} = V_w(\theta)$. Conversely dropping the advection term in (6.34), we recognize that the term $\partial_z[D_w \partial_z \theta]$ is analogous to the dispersion term introduced with (4.53). Hence, dispersion spreads θ it over a larger volume of soil, with the spreading rate described by the diffusivity D_w, in much the same way as molecular diffusion spreads dissolved substances. Since the corresponding parameter functions vary over a wide range (Figure 6.17), transient water flow is a highly nonlinear phenomenon.

To recognize the manifestation of this nonlinearity, we consider some nonuniform distribution of the water content at time t. In a region with a higher value of θ, the velocity is higher and the water will overtake the water in the region ahead if θ is smaller there. Advection thus leads to self-sharpening fronts and, conversely, to tails that are spread out even further (Figure 6.18). While advection alone would thus create a shock – a moving discontinuity in θ – dispersion smooths the front. We may suspect that for a constant infiltration flux, these counteracting effects lead to a constant shape of the advancing front since they are balancing each other locally. Indeed, this was proved by *Philip* [1957*b*] for the long-time limit.

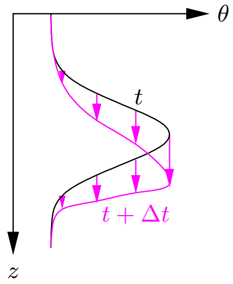


Figure 6.18.
Sketch for effect of pure advection with θ-dependent velocity on shape of infiltrating pulse.

Of course, dispersion also leads to a further spreading of the tail. Compared to that caused by the nonlinear advection, this effect is negligible, however. Focussing again on the advancing front and recalling Section 4.1.3, we may introduce the length scale

$$\lambda_w(\theta) := \frac{D_w(\theta)}{V_w(\theta)} \tag{6.37}$$

to characterize the typical extent of the front. In analogy to 4.38, we call this quantity the *dispersivity*. For transport distances smaller than λ, dispersion dominates advection and vice versa. As illustrated in the rightmost graph of Figure 6.17, $\lambda_w(\theta)$ ranges between 0.1 and 0.7 m for the sand and between some 0.3 m and practically infinity for the silt. Since advection and dispersion counteract each other in a progressing front, we expect some 0.1 m for the width of an infiltration front in the sand for a wide range of θ. On the other hand, the corresponding width in the silt will depend strongly on θ and it will be very much wider. Considering the tail, we first notice that both, advection and dispersion, act in the same way, namely as spreading. Hence, a lower bound for the spreading of the tail is obtained from the advective term alone. To this end, we consider two points with infinitesimal separation $\mathrm{d}z$, define the steepness $s = \mathrm{d}z/\mathrm{d}\theta$ of the tail, and find $s(t) = s(0) + t\,\mathrm{d}V_w(\theta)/\mathrm{d}\theta$. Thus, $\mathrm{d}V_w(\theta)/\mathrm{d}\theta$ is a lower bound for the rate of change of the tail's steepness.

The above qualitative discussion is corroborated by numerical simulations of a single infiltration pulse into dry uniform soil that will be discussed in Section 6.3.3 below. Before looking into the combined action of infiltration and drainage, we will study the partial processes individually.

6.3.1
Infiltration

We continue to consider the sand and the silt soil with Mualem-van Genuchten parameters given in Table 6.1, with a horizontal surface, horizontal layers, and a constant water table at depth $z = 0$. For times $t < 0$, the water phase is in static equilibrium with the water table, i.e., $j_w = 0$ in the entire profile

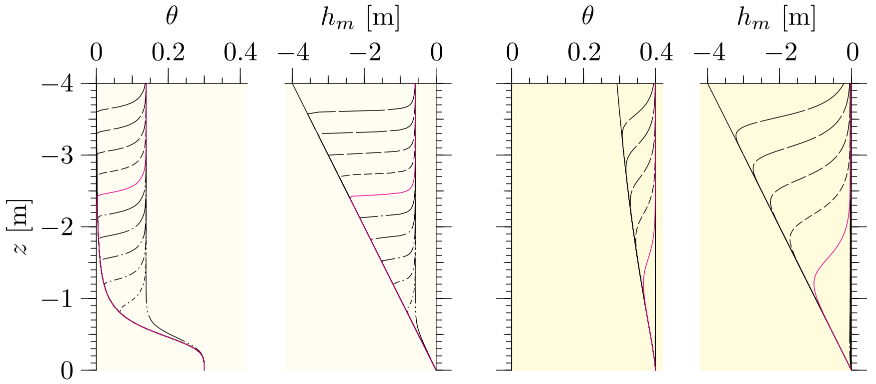


Figure 6.19. Infiltration into uniform sand (**left**) and silt (**right**). For $t < 0$, the water phase is in static equilibrium with the constant water table at 4 m depth. At $t = 0$, a constant infiltration with $j_w^0 = 5.56 \cdot 10^{-6}$ m s^{-1} (20 mm h^{-1}) sets in. The profiles are shown for $t = 0,\ 2,\ 4, \ldots$ h with those for $t = 10$ h drawn in magenta.

and $h_m = z$. Finally assuming a uniform infiltration flux, a one-dimensional representation suffices. As a passing remark, notice that there are situations in which the flow turns three-dimensional in spite of the fact that all the above conditions are satisfied. We will study such processes in Section 6.5 below.

Uniform Soil Let the groundwater table be 4 m below the soil surface where, for $t < 0$, $h_m = -4$ m. For $t \geq 0$, the water flux through the upper boundary is set to $j_w^0 = 5.56 \cdot 10^{-6}$ m s^{-1} (Figure 6.19). This corresponds to 20 mm h^{-1}, a heavy rainfall. For the sand, j_w is by a factor of about 20 smaller than the saturated conductivity K_0. For the silt, the corresponding factor is less than 2. In the regime of gravity flow, far behind the infiltration front and far above the water table, the matric head h_m adjusts such that $K(h_m) = j_w$. For the sand, this leads to $h_m = -0.58$ m and to $\theta = 0.14$, with the latter calculated with the van Genuchten parameterization. The corresponding values for the silt are $h_m = -0.032$ m and $\theta = 0.40$.

In the sand, the influence of the water table is essentially limited to the lowest 1.5 m. Towards the soil surface, the initial water content is below 0.01. As expected, the infiltration front quickly develops its invariant shape and advances with a constant velocity which is determined by the infiltrating flux and by the required increase $\Delta\theta$ of the water content. Only considering the gravity flow regime, which is an excellent approximation behind the front, and neglecting the initial water content yields $\Delta\theta = 0.14$ which leads to the velocity $j_w^0/\Delta\theta = 0.14$ m h^{-1}. For $t = 10$ h – the profiles distinguished in Figure 6.19 – this gives for the front an estimated depth of 1.4 m, which is in reasonable agreement with the simulation. The somewhat deeper advance of the numerically simulated front originates from the initial water content which leads to a slightly higher velocity. We further notice that also the width of the front is of the order estimated from the dispersivity (Figure 6.17) and that this width increases as expected towards the water table as θ increases. As a conclusion, given the hydraulic properties, constant infiltration into a deep profile of a uniform, coarse-textured soil can be predicted quite easily without recourse to numerical simulations.

Conceptually, the situation is similar for the silt. However, since this material has a much finer texture and a correspondingly broader distribution of pore-sizes, the influence of the water table extends over the entire profile. The infiltrating front is thus in constant transition and does not attain an invariant shape. We notice that the front advances considerably faster in the silt than in the sand. At first, this may appear counter-intuitive, since sand at this flux is a better conductor. However, the speed of the infiltration front is not determined by the conductivity which, in an unsaturated medium, will adjust itself such that the flux can be maintained. The velocity is determined by $\Delta\theta$ and this is clearly smaller in the silt. As a passing remark, the speed of the infiltration front does not necessarily equal the speed with which chemicals that are dissolved in the infiltrating water would be transported. Solute transport hinges on the question whether the infiltrating water bypasses or extrudes the resident water. Richards' equation, as a continuum model that does not contain the pertinent details of the pore space anymore, is useless for answering this question. Experimental evidence actually shows that in nature, the whole spectrum from almost complete bypassing to complete extrusion occurs. This aspect will be studied in more detail in Section 7.1.5.

Two-Layer Soil We revisit the two arrangements of layered soils from Section 6.2.2 but now with thicker layers which allows us to separated the impact of the layer interface from that of the water table. At time $t = 0$, the water phase is again in static equilibrium with the water table at $z = 0$. We consider two infiltration regimes (Figure 6.20), a low flux of $1.818 \cdot 10^{-7}\ \mathrm{m\,s^{-1}}$ ($0.65\ \mathrm{mm\,h^{-1}}$) and a high flux of $5.56 \cdot 10^{-6}\ \mathrm{m\,s^{-1}}$ ($20\ \mathrm{mm\,h^{-1}}$).

The low flux is two orders of magnitude below the effective saturated conductivity of the two-layer medium. For gravity flow, this leads to $h_m = -2.26$ m in the silt and to -0.896 m in the sand. Hence, silt for this flux is a better conductor than sand. Conversely, the high flux is near the saturated conductivity of the silt and sand is the better conductor. This is manifest in the form of the matric head at the interface between the two materials as already discussed for the stationary infiltration in Section 6.2.2.

Considering the evolution of the infiltration fronts, we first notice that in the sand, the dispersivity λ_w as shown in Figure 6.17 is small, on the order of 0.1 m, and does not vary much in the relevant range of θ between 0.04 and 0.14. This leads to rather sharp fronts for the low as well as for the high flux regime. We further notice that these fronts develop in the sand whether it is the upper or the lower layer. This is not astonishing for the upper layer since here the water flux at the upper boundary instantaneously jumps from 0 to

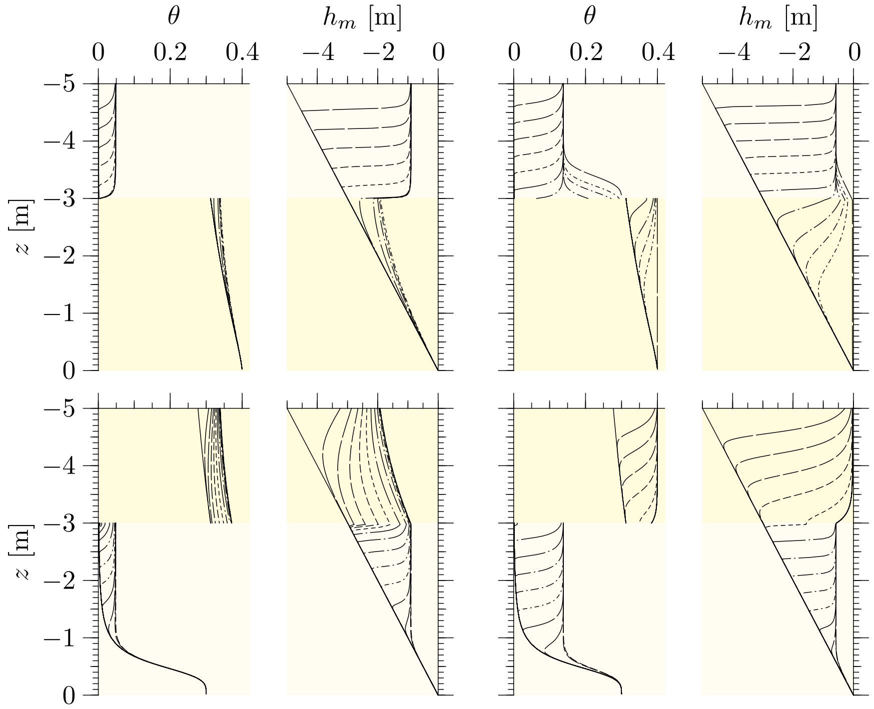


Figure 6.20. Infiltration into layered soil – sand over silt (**top**) and silt over sand (**bottom**) – with a constant water table at $z = 5$ m. For $t < 0$, the water phase is in static equilibrium. Constant infiltration starts at $t = 0$ with $1.818 \cdot 10^{-7}$ m s^{-1} (**left**) and $5.56 \cdot 10^{-6}$ m s^{-1} (**right**) corresponding to 0.65 mm h^{-1} and 20 mm h^{-1}. Profiles are shown for $t = 0,\ 1,\ 2, \ldots$ d and for $t = 0,\ 2,\ 4, \ldots$ h, respectively.

j_w^0 at $t = 0$. If sand is the lower layer, however, the water flux at its upper boundary may increase very gradually, as illustrated in the lower left graph of Figure 6.20. Still, given a sufficient layer thickness, a pronounced front develops. This demonstrates that the shape of a fully developed infiltration front is determined only by (i) the water flux, (ii) the hydraulic properties of the material, and (iii) the water content ahead of the front. In contrast, the details of how the eventual flux is attained, are irrelevant. The distance required for the front to develop is essentially the product of the front velocity in the sand and the time over which the flux at the upper boundary increases to its final value.

In the silt, on the other hand, λ_w is much larger than in the sand, for $\theta = 0.27$ considerably larger than 1 m, and it furthermore varies by about an order of magnitude for θ between 0.27 and 0.40. Consequently, an infiltration front develops for the high flux, but not for the low flux, for which the layer thickness is too small.

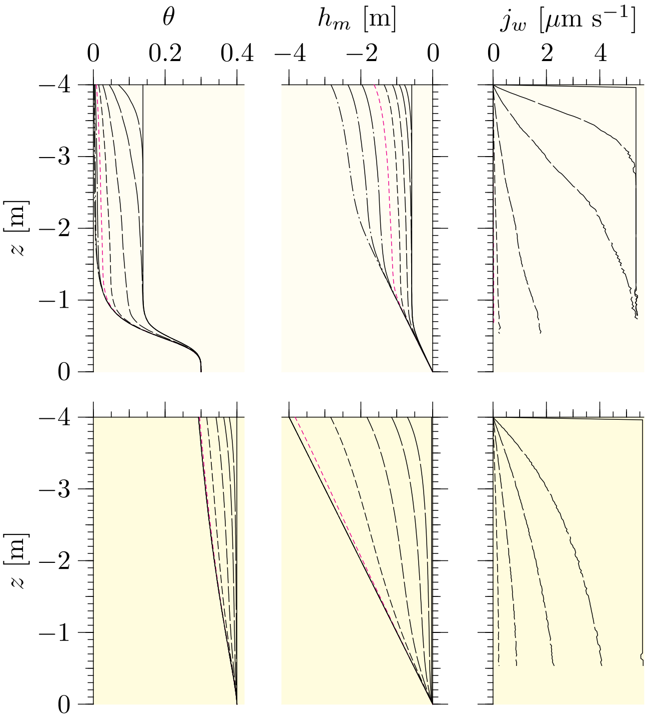


Figure 6.21. Drainage of sand (**top**) and silt (**bottom**) after the constant infiltration flux $j_w^0 = 5.56 \cdot 10^{-6}$ m s^{-1} (20 mm h^{-1}) ended abruptly at $t = 0$. Profiles of θ, h_m, and j_w are shown at $t = 0$, 1 h, 4 h, 16 h, 64 h,…,683 d. Notice the rapidly increasing time increment: the magenta lines correspond to some 43 days. Water flux is estimated from the numerical solution as $j_w = -K(\theta)[\mathrm{d}_z h_m - 1]$. Near the hydraulic equilibrium $\mathrm{d}_z h_m = 1$ this estimate becomes uncertain and is not plotted.

6.3.2
Drainage

We turn to the process complementary to infiltration, drainage, and consider a soil that is in dynamic equilibrium with the infiltration flux $j_w^0 > 0$ for $t < 0$. At $t = 0$, the infiltration stops abruptly and the soil drains under the influence of gravity, eventually reaching hydrostatic equilibrium with the constant water table at $z = 0$. We again consider the uniform soils sand and silt with parameters given in Table 6.1 and use a soil thickness of 4 m (Figure 6.21).

For $t < 0$, gravity flow is prevalent throughout most of the soil profile. With the interruption of flux, the transition occurs from $h_m \approx h_m^{\mathrm{grav}}$ to the hydrostatic equilibrium $h_m = z$. Apparently, and maybe somewhat unexpectedly, the silt drains very much faster than the sand. While after just one hour the water flux has dropped significantly throughout the entire profile in the silt, it remains practically unchanged in the sand for depths beyond some 1.2 m below the soil surface. Even in the silt, however, the transition to the static equilibrium is an exceedingly slow process. As we read from the h_m-profiles, it takes some 43 days for the silt to reach equilibrium while the sand is quite far away even after 2 years. The essential reason for this is that there are practically no pores in the sand that can retain water beyond about $h_m = -2$ m (see Figure 6.6). In the 4 m deep profile we consider here, this leads to a very low equilibrium water content in the upper 2 m, hence to a very low hydraulic conductivity. Since for drainage the gradient of the hydraulic head is limited to $-1 \leq \partial_z h_w \leq 0$, the drainage flux decreases very rapidly with θ. In addition, a much larger volume of water has to drain from the sand than from the silt.

We conclude from the very slow transition to the static equilibrium that in natural environments the vadose zone becomes rapidly decoupled from an underlying water table once a critical depth is exceeded. This critical depth depends on the soil water characteristic, specifically on the size of the smallest pores fraction, and on the typical time between rainfall events.

6.3.3
Infiltration Event

Turning towards more realistic scenarios, we consider a finite infiltration event, i.e., a constant infiltration for some limited time followed by drainage. Since we already know from the previous section that drainage is a very slow process, hence expect a very long tail for the pulse of water that will eventually propagate into the soil. We therefore look at a deep soil profile, choose a depth of 8 m, and furthermore assume that the water table is so deep that it does not affect the observed profile. This brings up the question of the appropriate initial condition. We take a rather pragmatic approach and, throughout the observed profile, choose a constant matric head h_m^0 such that a very small water content results.

Again employing the sand and the silt with parameters from Table 6.1 we realize that choosing the same initial condition for both material is not possible. Choosing h_m^0 too high, say $h_m^0 = -4$ m, turns the sand very dry – $\theta = 0.0006$ and $K(\theta) = 1.5 \cdot 10^{-13}$ $\mathrm{m\,s^{-1}}$ –, but the silt remains rather wet and conductive, $\theta = 0.35$ and $K(\theta) = 5.4 \cdot 10^{-8}$ $\mathrm{m\,s^{-1}}$. Choosing a value of h_m^0 that has the silt more dry, say $h_m^0 = -18$ m, turns the sand so dry and the conductivity so low that the numerical simulation gets very hard. We thus choose, somewhat arbitrarily, $h_m^0 = -4$ m for the sand and $h_m^0 = -18$ m for the silt.

As infiltration event, we consider a constant infiltration for 6 h during which 50 mm of water is applied to both soils. After the event, the flux through the upper boundary is zero (Figure 6.22). As expected, in the sand we find a rather sharp front for θ and h_m that widens only marginally as

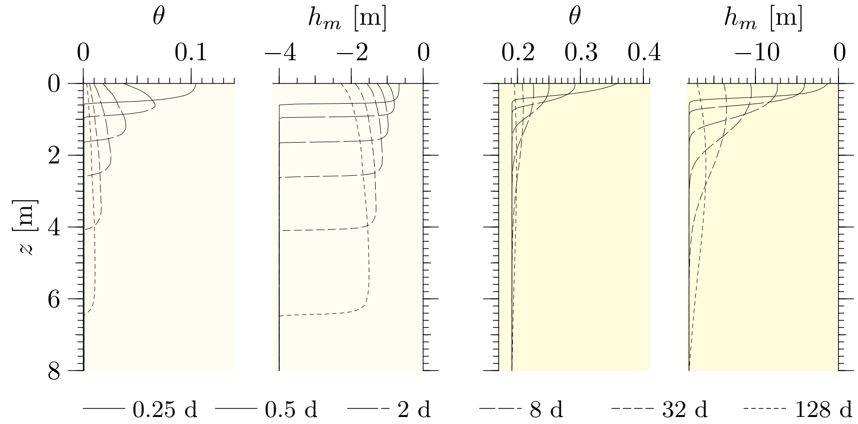


Figure 6.22. Propagation of an infiltration pulse – 50 mm of water within 6 hours – into uniform, initially dry soils. Initial matric head is $h_m = -4$ m for the sand (**left**) and $h_m = -18$ m for the silt (**right**). The solid lines show the profiles immediately after the end of the pulse at $t = 0.25$ d. For $t > 0.25$ d, the water flux through the upper boundary remained zero. Notice the different scaling of the horizontal axes for the two materials and the choice that $z = 0$ at the surface. The later is in contrast to other graphs in this chapter because the natural choice for $z = 0$ at the groundwater table is not possible, here.

the pulse progresses. Behind the front, h_m decreases very slowly both in space and time with a slightly stronger decrease near the soil surface. In contrast, the shape of θ behind the front changes rapidly during the initial stage. Looking at Figure 6.6 reveals that this originates in the transition through the strongly nonlinear part of the soil water characteristic. While $\theta(z)$ decreases rather rapidly for large values of θ, it becomes approximately constant for sufficiently low values due to the sheer steep increase of $V_w(\theta)$. In the silt, on the other hand, the self-sharpening of the front is hardly perceptible due to the overwhelming effect of dispersivity λ_w which increases strongly as θ decreases.

We notice in passing that simulating an infiltrating pulse calls for a representation of hysteresis since water content increases ahead of the front and decreases behind it. Hence a typical point in the soil profile would first move up the main wetting branch and then descend on some scanning drying branch. Since the hysteresis of $K(\theta)$ is negligible, while that of $\theta(h_m)$ is pronounced, the switching of the branches would not affect $V_w(\theta)$ but $D_w(\theta)$ would initially be larger on the drying branch. For the stable flow we consider in this section, the effect of hysteresis is thus only quantitative. As we will find in Section 6.5.2, however, hysteresis becomes a dominating aspect for unstable flow.

6.3.4 Evapotranspiration

In most environments, the dominant path for the loss of water from the vadose zone is transpiration and evaporation. *Transpiration* refers to water uptake by plant roots, transport through the plant tissue to the leafs, and evaporation from there. *Evaporation* on the other hand refers to direct vapor loss from the soil surface. Since the two processes are difficult to separate in nature, they are often aggregated into one, which is then called *evapotranspiration*. Such a separation is simple in a numerical simulation, however, and we consider in the following a water flux that leaves the soil either through evaporation or through transpiration.

Evaporation Again consider the uniform sand and the silt with a constant water table at $z = 0$ and constant infiltration flux $j_w^0 = 5.56 \cdot 10^{-6}$ $\mathrm{m\,s^{-1}}$. At $t = 0$, the infiltration stops abruptly and switches to an evaporation with $j_w^0 = -1.39 \cdot 10^{-8}$ $\mathrm{m\,s^{-1}}$, corresponding to -1.2 $\mathrm{mm\,d^{-1}}$ (Figure 6.23).

First consider the sand and notice that in addition to the drainage already observed in the previous section, a drying front develops in a thin surface layer where the matric head drops very rapidly. Indeed, by day 4, h_m at the surface already reaches -20 m, a prescribed lower bound which is imposed to prevent a run-away to $-\infty$. The evaporation flux cannot be maintained after day 4. The situation is very different for the silt. The profile equilibrates rapidly, within some 10 days, and the required evaporation flux is sustained easily by a minimal drop of h_m relative to its static equilibrium value, from -4.0 m to -4.3 m.

Transpiration To describe water uptake by plant roots, the Richards equation (6.2) must be enhanced by the volume extraction term γ_w which describes the volume of water that is extracted per unit volume of soil per unit time. The dimension of γ_w is thus T^{-1}. The enhanced Richards equation then becomes

$$\partial_t \theta - \nabla \cdot \left[\mathsf{K}_w(\theta)[\nabla h_m - \widehat{\mathbf{z}}]\right] = -\gamma_w \ , \qquad (6.38)$$

where $\widehat{\mathbf{z}}$ is the unit downward pointing vector. The inconspicuous extraction term γ_w actually hides a crucial and involved aspect, namely the distribution of plant roots, their physiologically determined activity, and their growth and decay. While a correct description of these aspects is crucial for modeling the coupling between soil, vegetation, and atmosphere, a most simple model suffices for our current interest. We first assume that the plant roots are able to take up all the water required by the prescribed flux j_w^0 and that the distribution of roots does not change in time. Following *Šimůnek et al.* [1994], we then assume that the roots' activity only depends on the matric potential in the surrounding soil. Hence, the problem reduces to describing how the extraction of the total transpiration flux j_w^0 is distributed in the soil. This is assumed to be proportional to the product $f(\mathbf{x})\alpha(h_m)$, where $f(\mathbf{x})$ is the density function of the root distribution and $\alpha(h_m)$ is the activity function

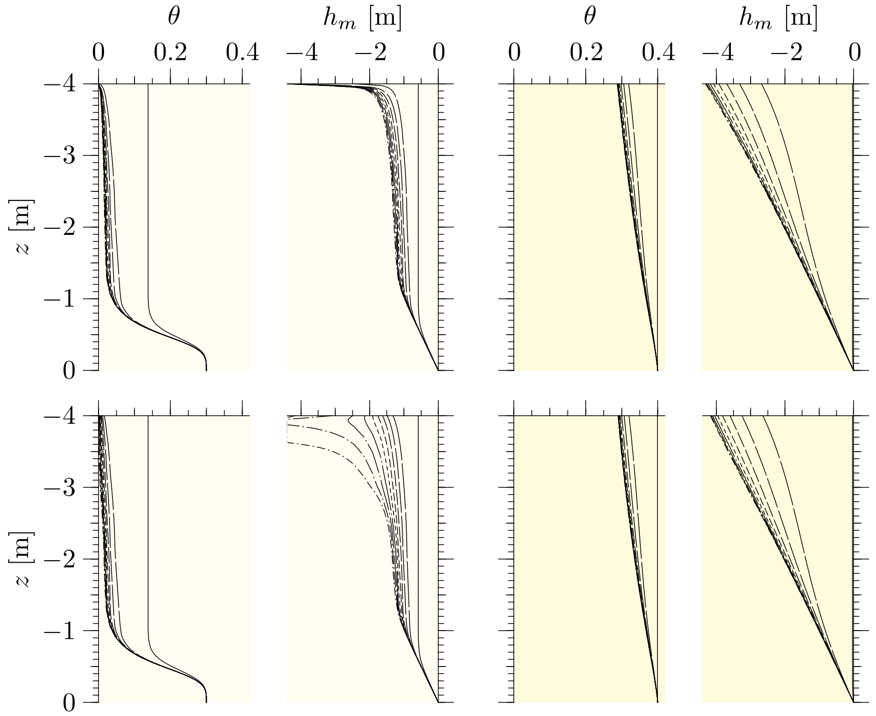


Figure 6.23. Evaporation (**top**) and transpiration (**bottom**) from uniform sand (**left**) and silt (**right**) with constant water table at $z = 4$ m. At $t = 0$, the water flux through the upper boundary jumps from $5.56 \cdot 10^{-6}$ m s^{-1} (20 mm h^{-1}) to $-1.39 \cdot 10^{-8}$ m s^{-1} (-1.2 mm d^{-1}). Profiles of θ and h_m are shown for $t = 0$, 2 d, 4 d,…, 18 d. Notice that, in contrast to Figure 6.21, time increments are constant.

of the roots. The function $\alpha(h_m)$ accounts for the fact that water uptake is only possible within a certain range of h_m which depends on the plant species. Most plant roots require oxygen for an optimal functioning hence there is an upper bound for h_m. On the other hand, if h_m in the soil is too low, uptake is impaired by the limited ability of the plant to reduce its potential and the resulting decrease, or even reversion, of the hydraulic gradient. Further dependencies are often incorporated into α, for instance a dependence on j_w^0 which would account for the resistance of the plant tissue to the water flux and the corresponding increase of the potential in the root.

Denoting the flow domain of interest by Ω and the area of its interface with the atmosphere with A, the extraction term may then be written as

$$\gamma_w(\mathbf{x}) = \frac{f(\mathbf{x})\alpha\big(h_m(\mathbf{x})\big)}{\int_\Omega f(\mathbf{x}')\alpha\big(h_m(\mathbf{x}')\big)\,\mathrm{d}\mathbf{x}'} j_w^0 A \; . \tag{6.39}$$

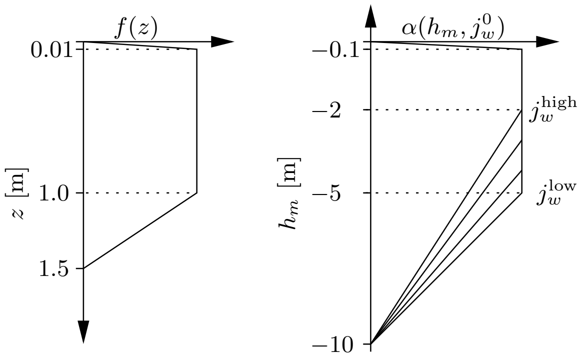


Figure 6.24. Root distribution density $f(z)$ and root activity $\alpha(h_m, j_w^0)$ for simple root model used in simulation of transpiration. No active roots are present right at the surface. The dependence of α on j_w^0 is motivated by resistive losses along the flow path in the plant and is also linear.

which, as required, yields $\int_\Omega \gamma_w(\mathbf{x}')\,\mathrm{d}\mathbf{x}' = j_w^0 A$. For the following simulations, we consider a one-dimensional situation and use for $f(z)$ and $\alpha(h_m)$ the piece-wise linear representations shown in Figure 6.24.

We consider the same situation as above for evaporation with the sole difference that the flux $j_w^0 = -1.39 \cdot 10^{-8}$ m s^{-1} (-1.2 mm d^{-1}) is not extracted at the surface any more but throughout the root zone. Profiles of θ and h_m are shown in the lower part of Figure 6.23.

For the silt, the differences between pure evaporation and pure transpiration are hardly perceptible. The water content, and with it the conductivity, is so high that very small deviations from static equilibrium suffice to maintain the required flux, wherever in the profile it is extracted.

For the sand, however, the situation is different. Since the water need not raise to the surface anymore, but only to the root zone, the flux j_w^0 can be maintained much longer and h_m only drops below -20 m by day 18. This illustrates how a deep root zone acts as a buffer that can supply vegetation with water through prolonged periods of drought even in coarse-textured soils. It also illustrates that soil beneath vegetation is typically much dryer than under a bare surface. This fact usually escapes superficial observation because vegetation does tend to make the top few millimeters slightly wetter as it better retains water vapor, which then condenses during cooler nights.

6.3.5
Natural Atmospheric Forcing

The forcing of the vadose zone by the atmosphere consists of sporadic rain events that are separated by extended dry periods where evapotranspiration prevails. While this general pattern is found almost everywhere, the amounts of water involved in the different processes varies greatly with climate zone, as does the duration of the infiltration and evaporation periods. In tropical regions, a diurnal cycle predominates with heavy showers in the early afternoon while in monsoon regions the annual cycle is dominant with a rainy season of few months duration and rather dry weather for the remainder of the year. Intermediate between these extreme are the humid climates where rain showers occur throughout the year with irregular intensities and temporal separations but with some clustering of events with durations on the order of weeks. A more intermittent variant of this regime is found in semi-arid and arid regions where short rain event with durations of hours to at most some days are separated by dry periods of many months to several years. Combining the range of climate zones with the range of soil types and adding in the vegetation and anthropogenic landuse – where we notice that these factors are not independent – leads to a diverse spectrum of phenomena.

As a generic example, we consider two uniform profiles, sand and silt, again with a constant water table at $z = 8$ m, deep enough to facilitate the separation of atmospheric forcing from the influence of the groundwater. The flow is driven with a measured time series of natural rainfall, negligible evaporation, and an assumed constant transpiration of -1 $\mathrm{mm\,d^{-1}}$ corresponding to a perennial gras cover. The time series stems from an experimental site near Bülach, Switzerland, and encompasses 200 d, starting in early May, 1991. Mean precipitation during this period was 3.11 $\mathrm{mm\,d^{-1}}$ with a maximum daily mean of 69 mm. For the distribution of the transpiration flux, the simple root model shown in Figure 6.24 was employed. To initialize the soil water state, h_m is set to the value that corresponds to gravity flow with the net infiltration flux $j_w^0 = 2.11$ $\mathrm{mm\,d^{-1}}$. This corresponds to -1.12 m for the sand and to -5.61 m for the silt. The simulation is then run for 40 d with a constant precipitation of 3.11 $\mathrm{mm\,d^{-1}}$ until $t = 0$. Such an initialization into some mean state is crucial since, depending on the forcing, the relaxation time of the soil may be very long. Starting from static equilibrium, for instance, the abnormal initial condition would be manifest in the sand for several months. The evolution of $\theta(z)$ and $h_m(z)$ after initialization, for $t > 0$, is shown in Figure 6.25. Profiles of θ and h_m for the most intense rain event during days 141. . . 144 are given in Figure 6.26. During this event, 109 mm of water infiltrated into the soil.

The dynamics of the water phase in the sand is dominated by the sharp infiltration fronts that develop from the individual rain events. Already small isolated events, like for instance during days 4. . . 7 where a total of 8 mm of water infiltrated, typically advance right to the water table even though the amplitude is very small. Larger groups of small events merge and eventually form one single pulse. An example is the period between days 76 and 95 where on average just 2.95 $\mathrm{mm\,d^{-1}}$ infiltrated, slightly less than the average for the entire period, with a maximum of just 8 $\mathrm{mm\,d^{-1}}$. Such mergers result from the higher velocity in regions with higher values of θ which in turn reflects the increase of K with θ. A consequence of this is that larger pulses, or merged groups of individual pulses propagate faster than smaller ones. This can be seen by comparing the previous group with the one that originated between days 21 and 52 and infiltrated at total of 201 mm with