1

Introduction

Partial differential equations (PDEs) have become enormously successful as models of physical phenomena. With the rapid increase in computing power in recent years, such models have permeated virtually every physical and engineering problem. The phenomena modeled by partial differential equations become increasingly complicated, and so do the partial differential equations themselves. Often, one wishes a model to capture different aspects of a situation, for instance both convective transport and dispersive oscillations on a small scale. These different aspects of the model are then reflected in a partial differential equation, which may contain terms (operators) that are mathematically very different, making these models hard to analyze, both theoretically and numerically.

A computational scientist is therefore often faced with new and complex equations for which an efficient solution method must be developed. If one is lucky, the equation is of a well-known type, and it is fairly easy to find efficient methods that are simple to implement. In most cases, however, one is not so lucky; good methods may be hard to find, and even good methods may be hard to implement.

A strategy to deal with complicated problems is to "divide and conquer". In the context of equations of evolution type, a rather successful approach in this spirit has been *operator splitting*.

The idea behind this type of approach is that the overall evolution operator is formally written as a sum of evolution operators for each term (operator) in the model. In other words, one splits the model into a set of sub-equations, where each sub-equation is of a type for which simpler and more practical algorithms are available. The overall numerical method is then formed by picking an appropriate numerical scheme for each sub-equation and piecing the schemes together by operator splitting.

In an abstract way one can formulate the method as follows: We want to solve the Cauchy problem

$$\frac{dU}{dt} + \mathcal{A}(U) = 0, \qquad U(0) = U_0, \tag{1.1}$$

where $\mathcal{A}$ is some unspecified operator. Formally (but not very helpful from a computational point of view) the solution reads[1]

$$U(t) = e^{-t\mathcal{A}} U_0. \tag{1.2}$$

[1] Inspired by the case when $\mathcal{A}(U) = \mathcal{A}U$, where $\mathcal{A}$ is a finite matrix, we formally write $e^{-t\mathcal{A}}$ for the solution operator.

Assume that we can write $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ in some "natural" way, and that one can solve the sub-problems

$$\frac{dU}{dt} + \mathcal{A}_j(U) = 0, \qquad U(0) = U_0, \quad j = 1, 2, \tag{1.3}$$

more easily with formal solutions

$$U_j(t) = e^{-t\mathcal{A}_j} U_0, \quad j = 1, 2. \tag{1.4}$$

In its simplest form, operator splitting reads as follows: Let $t_n = n\Delta t$ (with Δt small and positive). Approximately, we hope that

$$U(t_{n+1}) \approx e^{-\Delta t \mathcal{A}_2} e^{-\Delta t \mathcal{A}_1} U(t_n). \tag{1.5}$$

For commuting operators we have that $e^{-tA_2} e^{-tA_1} = e^{-t\mathcal{A}}$ and the method would be exact. Taking it one step further, one could hope that

$$U(t) = e^{t\mathcal{A}} U_0 = \lim_{\Delta t \downarrow 0,\, t = n\Delta t} \left(e^{-\Delta t \mathcal{A}_2} e^{-\Delta t \mathcal{A}_1} \right)^n U_0, \tag{1.6}$$

(with a limit to be determined) which indeed is the celebrated *Lie–Trotter–Kato formula*. A numerical method is obtained if one replaces the exact solution operators $e^{t\mathcal{A}_j}$ by numerical approximations. All splitting methods are refinements of this basic set-up.

This approach may seem a bit primitive at first glance, but in fact operator splitting has several advantages. Since the operators in the new submodels may be very different, they may also require very different numerical and analytical techniques. Operator splitting allows one to exploit this, and the resulting numerical method may be both simpler to implement and more efficient. By operator splitting, one can combine specialized numerical methods that have been developed to solve a particular class of evolutionary problems (i.e., developed especially for one of the elementary operators) in a fairly straightforward manner. This way, one can choose from a toolbox of highly efficient and well-tested numerical methods for elementary operators that can be combined to solve complicated problems. Indeed, the operator-splitting framework offers great flexibility in replacing one scheme for an elementary operator with another scheme for the same operator. Moreover, the use of operator splitting may also reduce memory requirements, increase the stability range, and even provide methods that are unconditionally stable. For very high dimensional problems this may be the only feasible method. Finally, by resorting to operator splitting, it is also easy to add increasing complexity to a numerical model, since each new term can be an independent numerical module.

The idea of splitting sums of complicated operators into simpler operators that are treated separately, is both easy and fundamental, and as such has appeared under various names in different contexts. We will here indicate some of the historical development, with no ambition of providing a complete survey. One of the

first rigorous results is associated with the name of Trotter [268]. The fundamental question he asked was: Given two continuous semi-groups with corresponding generators, how can one define the semigroup corresponding to the sum of the two generators? This corresponds to the equations (1.1)–(1.6) above. The result in the case of finite-dimensional matrices goes back to Sophus Lie and important extensions were provided by Kato [152]; the result is often denoted as the *Lie–Trotter–Kato formula* or simply the *Trotter formula*. Applications by Trotter and Kato were to quantum mechanics. Several refinements of this method exist, for instance, the *Baker–Campbell–Hausdorff formula* expresses the operator $\tilde{A}$ with the property that $e^{-tA_2}e^{-tA_1} = e^{-t\tilde{A}}$.

In a more concrete setting (and prior to Trotter and Kato), Douglas, Peaceman, and Rachford [90, 220] introduced a method called the *alternating direction implicit* (ADI) scheme, where multi-dimensional problems were successfully reduced to repeated one-dimensional problems. The ADI method was soon applied to petroleum reservoir simulation. In the late 1960s, increased computer power made other methods viable for reservoir simulations. Starting in the late 1950s and early 60s, there was an extensive development in the Soviet Union, using what was coined *splitting methods* or the *fractional steps method* as a general method to study a large variety of problems in mathematical physics and several applications. Key advances were made by Yanenko, Samarskii, Marchuk and others. It is impossible here to survey the results obtained; rather we refer to Yanenko's monograph [278], and the comprehensive survey by Marchuk [203]. Related to these methods is the method of *locally one-dimensional* (LOD) methods, where a dimensional splitting in effect reduces the original problem to a series of one-dimensional problems. For a general survey of splitting methods we refer to Hundsdorfer and Verwer [129, Ch. IV]. For matrix-related methods we refer to [208]. Observe that one often finds the same method denoted by different names, and the same name used for different methods. This is due to wide applicability of the method, but it complicates an accurate historical description of the development.

Most of the refinements depend on further knowledge of properties of the underlying sub-problems. Detailed knowledge of the behavior of solutions can make rather powerful methods. Here we will analyze operator splitting for a class of nonlinear partial differential equations, see Section 1.2, with the property that the solutions are *rough*, i.e., the solutions are functions of limited regularity and may even contain jump discontinuities, called shocks, so that the equations have to be interpreted in the sense of distributions.

Operator splitting may not always be the right answer. The extent to which operator splitting will give an effective overall method depends on the coupling of different elementary operators and the dynamics of the evolution problem. If the elementary operators are weakly coupled—that is, if the interaction of the different physical phenomena has a long time scale—an operator-splitting scheme will be efficient over a wide range of sizes for the splitting steps. Furthermore, for higher-dimensional problems it may be the only feasible method. On the other hand, if

the operators interact significantly over a short time scale, operator splitting may be subject to severe restrictions on the splitting step. For nonlinear operators, interaction between elementary operators is often nonlinear, and splitting them into separate steps may result in large and unwanted errors. To prevent or remedy such splitting errors requires a thorough understanding of the underlying error mechanisms.

1.1 Purpose of the book

The purpose of this book is to give an introduction to various types of operator-splitting methods for constructing discontinuous, but physically relevant, solutions of nonlinear mixed hyperbolic-parabolic partial differential equations. The theory is illustrated by several examples and MATLAB code for most of the examples is posted on the web site

www.math.ntnu.no/operatorsplitting

The class of equations is very rich, and contains, for instance, hyperbolic conservation laws, heat (diffusion) equations, porous medium equations, two-phase reservoir flow equations, as well as (strongly) degenerate convection-diffusion equations with applications to sedimentation. These equations are frequently also referred to as degenerate (or degenerate parabolic) convection-diffusion equations. A significant part of this book is devoted to reporting the results of applying operator-splitting methods to a variety of convection dominated problems, including problems coming from flow in porous media, shallow water waves, and gas dynamics. Along the way we make an effort to provide enough (algorithmic) details so as to enable the readers themselves to implement the presented methods without too much effort. Another significant part of this book aims at introducing the reader to the basic parts of a theoretical foundation of operator-splitting methods for convection-dominated problems possessing solutions with limited regularity or even discontinuous solutions. Although the theory is restricted to problems consisting of scalar and weakly coupled systems of equations, it nevertheless provides guiding principles for designing accurate and efficient operator-splitting methods for systems of equations. A novelty of this book is that it develops a theoretical framework for operator-splitting methods based on recent 'hyperbolic' techniques. This enables us to treat the whole spectrum of equations in a unified manner, ranging from purely hyperbolic equations possessing shock wave (discontinuous) solutions, via degenerate parabolic equations admitting solutions with limited regularity or even shock wave solutions in the case of degeneracy on intervals, to uniformly parabolic convection-diffusion equations possessing smooth solutions. Furthermore, since it turns out that the hyperbolic arguments also apply to many weakly coupled systems of partial differential equations, we will in fact develop a convergence theory for a general class of weakly coupled systems of equations.

Hence this theory not only covers scalar equations but also many physically interesting weakly coupled systems of hyperbolic and parabolic equations. For some of the operator-splitting methods we identify intrinsic splitting-error mechanisms as well as present procedures for reducing the errors originating from these mechanisms.

1.2 The class of PDEs discussed in the book

We will in the following develop a theoretical framework for operator-splitting methods in the setting of systems of weakly coupled nonlinear partial differential equations of the type

$$\begin{aligned} u_t^\kappa + \sum_i F_i^\kappa(u^\kappa)_{x_i} &= \Delta A^\kappa(u^\kappa) + g^\kappa(U), \qquad (x,t) \in \mathbb{R}^d \times [0,T], \\ u^\kappa\big|_{t=0} &= u_0^\kappa, \quad \kappa = 1,\dots,K, \end{aligned} \tag{1.7}$$

with $U(x,t) = (u^1(x,t),\dots,u^K(x,t))$. The term *weakly coupled* means that the equations are coupled only through the source term $g^\kappa(U)$. The diffusive term is assumed to satisfy

$$\frac{dA^\kappa}{du}(u) \geq 0, \qquad \kappa = 1,\dots,K, \quad A^\kappa(0) = 0,$$

where the essential condition is the first one, under which (1.7) is referred to as *degenerate* or sometimes *degenerate parabolic*. A mild form of degeneracy occurs if for some κ we have $\frac{dA^\kappa}{du}(u) = 0$ for one or several values of u, in which case one often speaks of point degeneracy. A more severe form of degeneracy occurs if for some κ we have $\frac{dA^\kappa}{du}(u) = 0$ for u in some interval. In this case one often says that (1.7) is strongly degenerate. In other words, (1.7) is strongly degenerate if A^κ is constant on intervals. In general, the system (1.7) possesses solutions with limited regularity, i.e., weak solutions in the sense of distributions. Despite the restriction "weakly coupled", partial differential equations like (1.7) include several important model equations.

When $g^\kappa \equiv 0$ for all κ, the system (1.7) becomes a set of independent scalar partial differential equations. In particular, the scalar conservation law

$$u_t + \nabla \cdot f(u) = 0 \tag{1.8}$$

is a simple special case of (1.7) for $K = 1$. The regularized conservation law

$$u_t + \nabla \cdot f(u) = \Delta u \tag{1.9}$$

is another equation within the class analyzed here. Included is also the heat equation

$$u_t = \Delta u, \tag{1.10}$$

the one-point degenerate porous medium equation

$$u_t = \Delta u^m, \qquad m \geq 1, \tag{1.11}$$

the two-point degenerate two-phase reservoir flow equation

$$u_t + \Big(\frac{u^2}{u^2 + (1-u)^2}\Big)_x = \big(u(1-u)\big)_{xx}, \tag{1.12}$$

as well as the nonlinear, possibly strongly degenerate, convection-diffusion equation

$$u_t + \nabla \cdot f(u) = \Delta A(u), \qquad A' \geq 0. \tag{1.13}$$

An example of a strongly degenerate convection-diffusion equation is provided by the theory of sedimentation-consolidation processes [47]. In this theory a typical choice of A satisfies

$$A'(u) \begin{cases} = 0, & u \in [0, u_c], \\ > 0, & u \notin [0, u_c], \end{cases}$$

where $u_c > 0$ is a given constant, i.e., A is flat (constant) on the interval $[0, u_c]$. On $[0, u_c]$, equation (1.13) reduces to a hyperbolic equation (1.8). Consequently, degenerate convection-diffusion equations will in general possess all the features of hyperbolic conservation laws, including the existence of shock wave (discontinuous) solutions, the necessity of using weak solutions, the loss of uniqueness of weak solutions, the need for additional selection criteria (entropy conditions) to restore uniqueness, and so forth.

1.3 Operator splitting for initial-value problems

Let us revisit the abstract approach (1.1)–(1.6). Writing the system (1.7) as an abstract Cauchy problem

$$\frac{dU}{dt} + \mathcal{A}(U) = 0, \qquad U(0) = U_0, \tag{1.14}$$

with solution $U(t) = \mathcal{S}_t U_0$, the operator $\mathcal{A}$ can often be decomposed as a sum of elementary (simpler) operators in a natural way. As an example assume that $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$. Using the semigroup notation $U^j = \mathcal{S}_t^j U_0$ for the solution of

$$\frac{dU^j}{dt} + \mathcal{A}_j(U^j) = 0, \qquad U^j(0) = U_0, \qquad j = 1, 2, \tag{1.15}$$

we approximate the solution of (1.14) by

$$U(n\Delta t) \approx \big[\mathcal{S}^2_{\Delta t}\mathcal{S}^1_{\Delta t}\big]^n U_0. \tag{1.16}$$

An alternative splitting formula is obtained by reversing the order of the operators, but this will in general give a different approximation.

The aim is to prove a Trotter formula like

$$U(t) = \mathcal{S}_t U_0 = \lim_{n\to\infty} \left[\mathcal{S}^2_{\Delta t}\mathcal{S}^1_{\Delta t}\right]^n U_0 = \lim_{n\to\infty} \left[\mathcal{S}^2_{t/n}\mathcal{S}^1_{t/n}\right]^n U_0.$$

To obtain a numerical solution, we replace the exact solutions operators $\mathcal{S}_j$ by approximations, with the goal of proving that the Trotter formula still holds.

The operator splitting in (1.16) is only first-order accurate. As an alternative, one can use the so-called *Strang splitting*,

$$U(n\Delta t) \approx \left[\mathcal{S}^1_{\Delta t/2}\mathcal{S}^2_{\Delta t}\mathcal{S}^1_{\Delta t/2}\right]^n U_0. \tag{1.17}$$

which is formally second-order accurate for sufficiently smooth solutions. The two operator splittings (1.16) and (1.17) are examples of so-called *multiplicative* operator splittings, which will be the main focus in this book.

Multiplicative operator splitting is closely related to the ADI method. To explain the idea behind ADI, we replace the exact evolution operators in (1.15) by standard forward/backward Euler approximations. Writing $U^n = U(n\Delta t)$, the classical ADI method reads

$$\begin{aligned} U^{n+\frac{1}{2}} + \mathcal{A}_1\big(U^{n+\frac{1}{2}}\big) &= -\mathcal{A}_2\big(U^n\big), \\ U^{n+1} + \mathcal{A}_2\big(U^{n+1}\big) &= -\mathcal{A}_1\big(U^{n+\frac{1}{2}}\big). \end{aligned} \tag{1.18}$$

Another class of operator splitting is the so-called *additive operator splitting* (AOS). The first-order equivalent of (1.16) reads

$$U(n\Delta t) \approx \left[\frac{1}{2}(\mathcal{S}^2_{2\Delta t} + \mathcal{S}^1_{2\Delta t})\right]^n U_0, \tag{1.19}$$

whereas the second-order equivalent of the Strang splitting reads

$$U(n\Delta t) \approx \left[\tfrac{1}{2}(\mathcal{S}^1_{\Delta t}\mathcal{S}^2_{\Delta t} + \mathcal{S}^2_{\Delta t}\mathcal{S}^1_{\Delta t})\right]^n U_0. \tag{1.20}$$

There are two main motivations for these operator splittings. First of all, the result of an additive operator splitting is independent of the order in which the operators are applied. For the multiplicative operator splitting, the operators will generally not commute in the nonlinear case, which means that the result depends on the order in which the operators are applied. This is the main reason why AOS methods are gaining popularity within image processing, even though they generally are less accurate than multiplicative splittings. The second advantage of AOS methods is that, since the operators are applied independently, they can be computed in parallel. AOS methods are therefore often used in combination with parallel processing.

1.4 Operator splitting for convection-diffusion equations

As already mentioned, we will in this book study nonlinear evolutionary PDEs of mixed hyperbolic-parabolic type. By advocating operator splitting, the numerical solution of the abstract problem (1.14) is reduced to the numerical solution of simplified problems of the type (1.15), for which one may utilize highly efficient methods which are tailor-made for each simplified subproblem. In recent years, we have witnessed an immense activity in developing sophisticated numerical methods for hyperbolic partial differential equations. We refer to [15, 67, 101, 104, 107, 108, 115, 159, 175, 176, 260, 262] for an introduction to modern numerical methods for hyperbolic equations. It is a reasonable strategy to attempt to utilize some of these hyperbolic solvers as building blocks in numerical methods for convection-diffusion problems. Indeed, in this book we make use of a diversity of hyperbolic solvers, including monotone schemes such as the upwind and Godunov schemes, quasi-monotone schemes, front tracking, large-time-step Godunov or Glimm methods, characteristic Galerkin methods, second-order MUSCL schemes, and high-order nonoscillatory central schemes. It is well-known that an accurate numerical approximation of convective and diffusive processes is a very difficult matter. This is especially true if convection dominates diffusion, which is the quintessential case. Accurate numerical simulations in such cases are often complicated by excessive amounts of unphysical oscillations or numerical diffusion. Often numerical methods based operator splitting and modern hyperbolic solvers avoid undue amounts of oscillations and diffusion.

A typical splitting approach for convection-diffusion equations involves not only hyperbolic equations modeling convection effects, but also (possibly degenerate) parabolic equations imitating diffusion effects. In this book, we will rely on very simple finite-difference schemes to approximate these parabolic equations. However, there exists a diversity of numerical methods that have been developed over the last fifty years—including finite-difference, finite-volume, and finite-element methods. To learn about numerical methods for elliptic and parabolic equations we invite the reader to take a closer look at one or several of the references [40, 101, 104, 112, 129, 136, 154, 209, 230, 241, 247, 259–261, 269].

1.5 Rigorous analysis of operator-splitting methods

A key focus of this book is the analysis of what happens when the exact solution operators $\mathcal{S}_j$ are replaced by approximate solvers. To this end, we provide a general theoretical framework by which it follows that if the approximate solvers for $\mathcal{S}_j$, cf. (1.15), satisfy certain properties, then the corresponding operator-splitting method will converge to the exact solution of the underlying partial differential equation. This framework includes scalar and weakly coupled systems of nonlinear partial differential equations containing various combinations of hyperbolic and

parabolic effects, cf. (1.7). As a pedagogic device to convey to novice readers the fundamental parts of this framework, we will frequently illustrate the main methodological concepts and results on simplified problems.

Let us recall that operator-splitting methods can always be analyzed in terms of accuracy by straightforward Taylor expansions, at least formally. For recent work on such analysis of operator splitting from the point of view of the Lie operator formalism, see [171]. However, in terms of rigorous analysis, this approach is not satisfactory, since nonlinear partial differential equations in general will possess solutions that exhibit complex behavior in small regions in space and time, i.e., sharp transitions or even singularities like shock waves (discontinuities). Moreover, even if the underlying exact solution is smooth, it can be that the operator splitting is composed of solution operators $\mathcal{S}_j$ that may produce nonsmooth solutions. An example of this case is provided by viscous operator splitting of nonlinear convection-diffusion equations, in which the nonlinear convection operator may introduce discontinuities into an otherwise smooth solution.

The general convergence framework developed in this book, in the context of fully discrete operator-splitting methods for weakly coupled systems of equations containing a synthesis of hyperbolic and parabolic effects, is based on the so-called Kružkov L^1-entropy solution theory. This pioneering theory was originally developed by Kružkov [161] for first-order quasilinear hyperbolic equations and only recently extended by Carrillo [50] to second-order quasilinear degenerate parabolic equations. Our convergence theory includes and extends previous (L^1) convergence results for problem-specific operator-splitting methods. For weakly coupled systems of hyperbolic conservation laws we also provide abstract L^1-error estimates for dimensional-splitting methods. Hence, in order to verify convergence (or convergence rates), one only has to check whether *each* method satisfies certain assumptions, whereupon convergence follows. When applied in a specific situation, these abstract error estimates avoid Kružkov's usual doubling of variables. We consider a variety of semi-discrete and fully discrete operator-splitting methods, including dimensional splitting, viscous splitting, flux splitting, and source-term splitting, and verify that the conditions needed to apply the abstract convergence results hold. The main advantage of the L^1-approach is that it makes it possible to have a unifying convergence theory for hyperbolic, parabolic, and mixed hyperbolic-parabolic problems. In the parabolic setup, where alternative approaches are possible, the L^1-approach has the advantage that it yields results that are independent of the Peclet number, i.e., the ratio of convection forces to diffusion forces. At this point it should be stressed that the development of a unifying theoretical framework is possible only due to the recent forming of a mathematical theory for discontinuous solutions of strongly degenerate convection-diffusion equations [20, 45, 48–50, 55–58, 102, 147, 148, 178, 202, 205, 207, 222, 231, 232, 251, 252, 270].

The idea behind operator splitting is certainly an old one and has been comprehensively described, for example, in [63, 129, 203, 206, 246]. The new, and to a certain extent original, aspect of our presentation lies in the systematic

use of numerical schemes and mathematical theory associated with hyperbolic equations. We use this hyperbolic approach to construct splitting methods and a corresponding unifying convergence theory for degenerate convection-diffusion problems. In doing so we are building on and extending our previous (unified) analysis of operator splitting methods [119, 120], which again is based on ideas that have evolved over several years [34–37, 43, 93, 97, 100, 114, 121, 124–126, 131–133, 139–143, 145, 146, 182]. Consult [93] for a review of this activity. Another unconventional facet of our presentation is the focus on splitting methods that allow for large time-steps. The use of large time-step methods for the convection step, like front tracking, has some advantages. For example, in the setting of a nonlinear convection-diffusion equations and an implicit diffusion solver, the resulting operator-splitting methods become unconditionally stable in the sense that there is no CFL condition (named for its originators Courant, Friedrichs, and Lewy) restricting the time step. Indeed, it has always been our firm belief that the time-step in a numerical method should be dictated by the dynamics of the equation and not by the spatial discretization. For convection-diffusion equations, it turns out that a practicable time-step is highly dictated by the degree of (nonlinear) interaction between convective and diffusive forces. Unfortunately, large time-steps can lead to fronts (sharp transitions in the solution) that are too wide: A recurrent theme in this book is that it is possible to identify and reduce this kind of splitting errors, thereby yielding accurate large time-step methods, along the lines of the approach initiated in [146] and further developed and analyzed in [100, 140–143]. The approach in [146] was motivated by an idea introduced in [92] and further expanded on in a series of papers [74–78].

Besides viscous splitting methods for convection-diffusion problems, we will devote considerable attention to methods for hyperbolic problems based on dimensional splitting [71, 125, 139, 180, 182] as well as source splitting [170, 177, 223, 253, 254].

1.6 Topics not treated in the book

Before we end this introductory chapter, let us list some important topics that are not treated in this book. First of all, there are of course many numerical approaches that do not rely on operator splitting, cf. the lists of references given above for hyperbolic and parabolic problems and cf. [6, 32, 44, 46, 51, 60, 96, 98, 99, 102, 111, 147, 164, 207, 216, 217] for mixed hyperbolic-parabolic problems. Regarding convection-diffusion problems, we omit the Godunov-mixed operator splitting methods [81–84, 274] and the recent fast explicit operator-splitting methods [61, 62]. Next, we do not address the numerical solution of so-called elliptic-parabolic problems, see, e.g., [3, 5, 24, 130, 137, 138, 197, 219, 236, 237]. Moreover, we do not discuss hyperbolic and mixed hyperbolic-parabolic problems with discontinuous flux, see, for example, [149–151, 153, 266, 267]. Nor

do we go into operator splitting for reaction-diffusion equations, as in for example [86–88, 235, 245]. Nonlinear convection-diffusion equations can be seen as toy models for the fundamental equations of fluid flow—the Navier–Stokes equations—and viscous operator splitting for the Navier–Stokes equations has been analyzed and applied in a great number of papers, see [16–18, 200, 279–287] and, for example, [192, 193, 272] for parallel splitting methods. None of these papers will be examined herein. For operator splitting applied to the Boltzmann equation, which describes the statistical distribution of particles in a fluid, see, for example, [69]. Another important class of equations that is omitted in this book is that of the Hamilton–Jacobi and Hamilton–Jacobi–Bellman equations, see [14, 103, 131–133, 188, 189, 210, 244, 248, 263]. operator-splitting methods for such equations arising in the context of finance have been used and analyzed in several papers, see, for example, [12, 13, 42, 264, 265]. Let us also mention that a convergence theory for splitting methods in the setting of maximal monotone operators on Hilbert spaces has been developed in [188], see also [210]. For operator splitting of the KdV (Korteweg–de Vries) equation, which models waves on shallow water surfaces, we refer to [121, 123, 255] and the references therein. Operator splitting from the point of view of semigroups has been a topic of study in for example [89, 155–157]. Variants of the Schrödinger equation, which is the fundamental equation of nonrelativistic quantum mechanics, have been approximated by operator splitting in [10, 11, 27, 273], see also [127] for the Maxwell–Dirac system.

1.7 Organization of the book

The book is organized as follows: In Chapter 2 we give the reader a taste of the content in terms of some simple examples of elementary operator splittings. Moreover, we discuss the convergence of splitting approximations and briefly touch upon errors common to them. The splittings will all be semi-discrete in the sense that there will be analytical solutions available for the split-operators. In Chapter 3 we present central elements of the mathematical framework in which to analyze, both from a mathematical and numerical point of view, second-order quasilinear degenerate parabolic equations. This theory, which can be viewed as a generalization of the well-known Kružkov theory [161] of entropy solutions for first-order hyperbolic conservation laws, provides the foundation for the convergence theory for operator splitting developed in the last section of the chapter. The theory is demonstrated in Chapter 4 by applying it to one-dimensional convection-diffusion problems. We consider a variety of semi-discrete and fully discrete splitting methods and verify that the conditions needed to apply the abstract convergence theory hold. Moreover, we present several numerical examples to highlight the use of operator splitting as a basis for developing efficient numerical schemes for convection-diffusion equations. In particular, we discuss underlying error mechanisms, and in some cases suggest strategies to reduce the splitting errors. In Chapter 5 we

extend the approach originating in Chapter 3 to yield not only convergence of operator-splitting methods, but also precise error estimates, at least in the context of hyperbolic problems. We also present many numerical examples using dimensional splitting, discuss error mechanisms, and go into how to choose the splitting step to optimize runtime versus numerical errors. Chapter 6 is devoted to numerical examples for systems of equations; these systems are not covered by the rigorous analysis in the previous chapters. In particular, we discuss applications from porous media flow and for two systems of conservation laws: the Euler equations of gas dynamics and the shallow-water equations. Finally, the purpose of Appendix A is to provide the novice reader with a brief introduction to numerical methods for hyperbolic problems, many of which will be used as building blocks in the splitting algorithms discussed in the following chapters.

The theory is illustrated by many examples throughout the text. For many of the examples in the book, runtimes are given. Note that the examples were developed over a period of several years and on various computers. Thus the runtimes will be considerably lower today, however, the relative times should remain unchanged.

The purpose of computing
is insight, not numbers.
— R. W. Hamming

1.8 Matlab programs

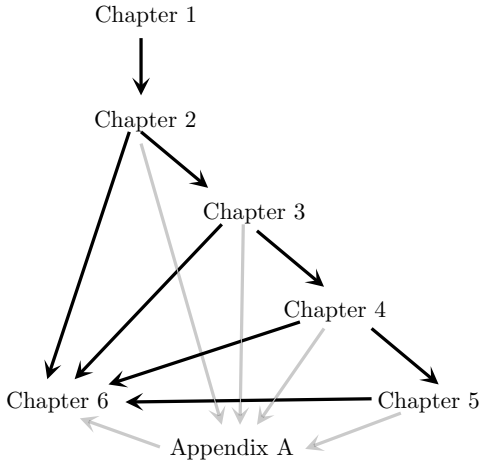
The theory described in this book is applicable in many different settings. One of the attractions is that it is reasonably easy to develop computer codes that can be used for the computation of (approximate) solutions which can be used for theoretical study as well as numerical results. To make the transition from theory to computer code easier, we offer computer codes in Matlab (version 7.7, R2008b) for almost all examples in the book. Note that the front tracking code is considerable slower when programmed in Matlab compared with codes in C or C++. Matlab codes are posted on the web site

www.math.ntnu.no/operatorsplitting

Feel free to use and modify them. Please let us know if you find bugs or possible improvements which we can post on the web site. If the computer codes are used for scientific work, we ask that you refer to the present book. We have decided to keep the computer codes on the web rather than include them in the book. The reasons for this are threefold: (i) It is easier to correct and update on the web than in a printed book; (ii) The book becomes smaller and can be used independently of the computer codes; (iii) Computer codes, and in particular their syntax and structure, change more rapidly than mathematical theory.

1.9 A guide to the reader

'Begin at the beginning,' the King said gravely,
'and go on till you come to the end: then stop.'
— Alice's Adventures in Wonderland



Chapter 6 focuses on applications of operator splitting in various contexts, and can be read independently of Chapters 3–5. The thin arrows mark a path through Appendix A for those unfamiliar with numerical methods for conservation laws.

1.10 Notation

For the d-dimensional ball with radius $r > 0$, we use the notation $\mathcal{B}_r$, i.e.,

$$\mathcal{B}_r = \left\{ x \in \mathbb{R}^d \;\middle|\; |x| \le r \right\}.$$

For partial derivatives we use all the different standard notation, e.g.,

$$\begin{aligned} \frac{\partial f}{\partial x_j}(x) &= f_{x_j}(x) = \partial_{x_j} f(x), \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d, \\ \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x) &= D^\alpha f(x), \quad \alpha \text{ multi-index.} \end{aligned}$$

Special differential operators are as usual given by

$$\begin{aligned} \Delta f(x) &= \sum_j \frac{\partial^2 f}{\partial x_j^2}(x), \\ \nabla f(x) &= (f_{x_1}(x), \dots, f_{x_d}(x)), \\ \nabla \cdot F(x) &= \sum_j \partial_{x_j} F_j(x), \quad F = (F_1, \dots, F_d). \end{aligned}$$

We will frequently be working in Lebesgue spaces, and we use standard notation: If $\Omega \subseteq \mathbb{R}^d$, we have for functions $f \colon \Omega \to \mathbb{R}$ that

$$\begin{aligned} \|f\|_{L^p(\Omega)} &= \begin{cases} \left(\int_\Omega |f(x)|^p \, dx \right)^{1/p}, & \text{for } p \in [1, \infty), \\ \operatorname{ess\,sup}_{x \in \Omega} |f(x)|, & \text{for } p = \infty, \end{cases} \\ L^p(\Omega) &= \{ f \colon \Omega \to \mathbb{R} \mid \|f\|_{L^p(\Omega)} < \infty \}. \end{aligned}$$

Local versions of the same spaces are defined by

$$L^p_{\mathrm{loc}}(\Omega) = \left\{ f \colon \Omega \to \mathbb{R} \;\middle|\; f\chi_K \in L^p(\Omega) \text{ for all compact sets } K \right\},$$

where we use the notation χ_K for the characteristic function of the set K. Let $C^p = C^p(\Omega)$, $p = 1, \dots, \infty$, denote the space of functions $f \colon \Omega \to \mathbb{R}$ possessing continuous partial derivatives of order $\le p$. In addition

$$C_0^p = C_0^p(\Omega) = \left\{ f \in C^p(\Omega) \;\middle|\; \operatorname{supp} f \text{ compact} \right\}. \tag{1.21}$$

For vector-valued functions $f \colon \Omega \to \mathbb{R}^K$ we write $C^p(\Omega; \mathbb{R}^K) = C^p(\Omega; \mathbb{R})^K$, etc., for the corresponding spaces.

More generally, for $\Pi_T = \mathbb{R}^d \times (0, T]$, we will often need to consider functions $u \colon \Pi_T \to \mathbb{R}^K$ as elements of various Bochner spaces; that is, we consider the functions $t \mapsto u(\,\cdot\,, t)$. For instance, we will employ the space

$$L^\infty(0, T; L^1(\mathbb{R}^d)) = \{ u \colon \Pi_T \to \mathbb{R} \mid \operatorname{ess\,sup}_{t \in (0,T]} \|u(\,\cdot\,, t)\|_1 < \infty \}.$$

The space $C(0,T;L^1(\mathbb{R}^d))$ consists of functions $u\colon \Pi_T \to \mathbb{R}$ such that the map $t \mapsto u(\,\cdot\,,t)$ is continuous in the L^1 norm. Finally, the space $C(0,T;L^1(\mathbb{R}^d;\mathbb{R}^K))$ consists of functions $u\colon \Pi_T \to \mathbb{R}^K$ with $t \mapsto u(\,\cdot\,,t) \in \mathbb{R}^K$ continuous in the norm in $L^1(\mathbb{R}^d;\mathbb{R}^K)$.

The Lipschitz constant of a function $f\colon \Omega \to \mathbb{R}^K$ is by definition

$$\|f\|_{\mathrm{Lip}} = \|f\|_{\mathrm{Lip}(\Omega)} = \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|}. \tag{1.22}$$

The corresponding space of Lipschitz functions is given by

$$\mathrm{Lip}(\Omega) = \left\{ f\colon \Omega \to \mathbb{R}^K \;\middle|\; \|f\|_{\mathrm{Lip}(\Omega)} < \infty \right\}, \tag{1.23}$$

with local version

$$\mathrm{Lip}_{\mathrm{loc}}(\Omega) = \left\{ f\colon \Omega \to \mathbb{R} \;\middle|\; \|f\|_{\mathrm{Lip}(K)} < \infty \text{ for each compact set } K \subseteq \Omega \right\}. \tag{1.24}$$

We will need the concept of total variation for a function, which is defined as follows: Consider first the one-dimensional case. For $f\colon [a,b] \to \mathbb{R}$ ($a = -b = -\infty$ permitted) we let

$$\mathrm{T.V.}\,(f)_{[a,b]} = \mathrm{T.V.}\,(f) = \sup_{a<x_0<\cdots<x_n<b} \sum_j |f(x_{j+1}) - f(x_j)|\,, \tag{1.25}$$

where the supremum is over all finite partitions $x_0 < x_1 < \cdots < x_n$. For functions in L^p spaces a refinement is needed (often called *essential variation*): We still use the definition above, but restrict the points $x_0 < x_1 < \cdots < x_n$ to points of approximate continuity of f, thereby obtaining a definition that is independent of the equivalence classes used in the proper definition of L^p spaces. Functions of bounded variation are defined as follows:

$$\mathrm{BV}([a,b]) = \{f \in L^1([a,b]) \mid \mathrm{T.V.}\,(f) < \infty\}. \tag{1.26}$$

For functions of several variables we use the following definition: Let $f\colon \mathbb{R}^2 \to \mathbb{R}$. Then we define the Tonelli variation

$$\mathrm{T.V.}\,(f) = \int_{\mathbb{R}} \mathrm{T.V.}\,(f(\,\cdot\,,y))_x \, dy + \int_{\mathbb{R}} \mathrm{T.V.}\,(f(x,\,\cdot\,))_y \, dx, \tag{1.27}$$

where $\mathrm{T.V.}\,(\,\cdot\,)_x$ and $\mathrm{T.V.}\,(\,\cdot\,)_y$ denote total variation with respect to the x and y variables respectively. Extensions to n variables are straightforward.

A function $f \in L^1(\Omega)$ is said to be of bounded total variation if its first-order derivative in the sense of distributions can be represented by a finite Radon measure, more precisely,

$$-\int_\Omega f \frac{\partial \phi}{\partial y_j}\, dy = \int_\Omega \phi \, d\mu_j, \qquad \phi \in C_0^\infty(\Omega),\ j = 1,\dots,d \tag{1.28}$$

with $|\mu_j|(\Omega) < \infty$. The set of all functions of bounded total variation is denoted by $\mathrm{BV}(\Omega)$. We denote the total variation of f by $|Df|_\Omega$ and define it by

$$|Df|_\Omega = \sup\left\{ \int_\Omega f \, \nabla \cdot \phi \, dy \mid \phi \in C_0^\infty(\Omega; \mathbb{R}^n), \|\phi\|_\infty \le 1 \right\}. \tag{1.29}$$

The local version is defined as follows: We say that a function $f \in L^1_{\mathrm{loc}}(\Omega)$ is in $f \in \mathrm{BV}_{\mathrm{loc}}(\Omega)$ if for each open set V, whose closure is contained in Ω, we have $|Df|_V < \infty$. We equip $\mathrm{BV}(\Omega)$ with the norm

$$\|f\|_{\mathrm{BV}} = \|f\|_{L^1(\Omega)} + |Df|_\Omega \,, \tag{1.30}$$

which makes $\mathrm{BV}(\Omega)$ into a Banach space [4, p. 121]. Then using Riesz's theorem on functionals in the space of continuous functions, we obtain that $\mathrm{BV}(\Omega)$ can equivalently be defined as

$$\mathrm{BV} = \mathrm{BV}(\Omega) = \left\{ f \in L^1(\Omega) \;\middle|\; \|f\|_{\mathrm{BV}\Omega} < \infty \right\}.$$

Note that for a function $f \in L^1_{\mathrm{loc}}(\Omega)$, we have that $\mathrm{BV}_{\mathrm{loc}}(\Omega)$ if and only if

$$\int_{R^{d-1}} \mathrm{T.V.}\,(f(\tilde{x}))_{\hat{x}_i} \, d\tilde{x} < \infty \tag{1.31}$$

for all compact rectangles $R^{d-1} \subset \mathbb{R}^{d-1}$ where we have for each i that $x = (x_1, \ldots, x_d) = (x_1, \ldots, x_{i-1}, \hat{x}_i, x_{i+1}, \ldots, x_d)$, $\tilde{x} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)$ (see [288, Thm. 5.3.5]). It is well known that the following inclusions hold:

$$\mathrm{BV}(\Omega) \subset L^{\frac{d}{d-1}}(\Omega) \text{ for } d > 1 \text{ and } \mathrm{BV}(\Omega) \subset L^\infty(\Omega) \text{ for } d = 1.$$

Furthermore,

$$\mathrm{BV}(\Omega) \text{ is compactly imbedded into } L^p(\Omega) \text{ for } 1 \le p < \frac{d}{d-1}.$$

See, e.g., [4, 95, 288] for an extensive discussion about BV functions.

When discussing difference schemes, we shall also be needing discrete versions of these norms. For a sequence $U = \{U_i\}_{i \in \mathbb{Z}}$, we define

$$\|U\|_p = \begin{cases} \left(\sum_{i \in \mathbb{Z}} |U_i|^p\right)^{1/p} & \text{if } p < \infty, \\ \sup_{i \in \mathbb{Z}} |U_i| & \text{for } p = \infty. \end{cases}$$

This is extended in the natural way to several dimensions, if we let $i \in \mathbb{Z}^d$ denote a multiindex $i = (i_1, \ldots, i_d)$, and let the above sum and supremum be taken over $\mathbb{Z}^d$.

Throughout this book, by the notation Const_X we shall mean a "constant" depending on X only.