Studies in Systems, Decision and Control 539

Gennady Yu. Kulikov
Maria V. Kulikova

State Estimation for Nonlinear Continuous–Discrete Stochastic Systems

Numerical Aspects and Implementation Issues

Springer

Studies in Systems, Decision and Control

Volume 539

The series "Studies in Systems, Decision and Control" (SSDC) covers both new developments and advances, as well as the state of the art, in the various areas of broadly perceived systems, decision making and control–quickly, up to date and with a high quality. The intent is to cover the theory, applications, and perspectives on the state of the art and future developments relevant to systems, decision making, control, complex processes and related areas, as embedded in the fields of engineering, computer science, physics, economics, social and life sciences, as well as the paradigms and methodologies behind them. The series contains monographs, textbooks, lecture notes and edited volumes in systems, decision making and control spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the worldwide distribution and exposure which enable both a wide and rapid dissemination of research output.

Indexed by SCOPUS, DBLP, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

Gennady Yu. Kulikov · Maria V. Kulikova

State Estimation for Nonlinear Continuous–Discrete Stochastic Systems

Numerical Aspects and Implementation Issues

Springer

Gennady Yu. Kulikov
CEMAT, Instituto Superior Técnico
Universidade de Lisboa
Lisbon, Portugal

Maria V. Kulikova
CEMAT, Instituto Superior Técnico
Universidade de Lisboa
Lisbon, Portugal

ISSN 2198-4182 ISSN 2198-4190 (electronic)
Studies in Systems, Decision and Control
ISBN 978-3-031-61370-8 ISBN 978-3-031-61371-5 (eBook)
https://doi.org/10.1007/978-3-031-61371-5

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The purpose of this book is to address rapid progress in mathematical modeling and simulation, which intends to employ complex and stochastic models for making more accurate predictions, for formulating better strategies and for designing sophisticated policies. Several optimal filtering algorithms have been designed for treating stochastic systems, where the optimality is understood in a statistical sense. The notion of optimal filtering refers to the family of computational techniques that are utilized for estimating the state of time-varying models which are indirectly observed through noisy discrete-time measurements. The state of a system is considered to be a vector of random variables related to positions, velocities, orientations of the plant under investigation and so on, which describe entirely the stochastic system in use. This contemporary mathematical modeling and simulation trend entails the necessity for exploring and exploiting nonlinear and continuous-time stochastic models with potentially stiff behavior, for which the state estimation task becomes extremely nontrivial and complicated, allowing for a numerical solution, only. Therefore, computational algorithms, such as methods for the numerical solution of integrals and ordinary differential equations, as well as numerous optimization means, can be interpreted as state estimation rules. Eventually, the accuracy and efficiency of nonlinear filtering rely basically on various aspects of the implemented numerical integration and optimization schemes and approximations.

The traditional filtering solution to nonlinear stochastic systems is presented by the well-known and popular extended Kalman filter. Recently, two more advanced and accurate state estimation tools termed the cubature and unscented Kalman filters have been designed. Furthermore, these have given rise to the concept of universal Gaussian filtering with deterministically sampled expectation and covariance, which unifies and covers the majority of modern methods in filtering theory. The main topic of our study is why and how effective adaptive numerical integration means with automatic local and global error regulations contribute to the task of accurate state estimation in nonlinear continuous-time state-space stochastic models with additive noise and discrete-time measurements, including those whose dynamic behavior is stiff, within the framework of Gaussian filtering techniques. This discussion focuses mostly not on statistical properties of determining posterior densities of the stochastic

model's state but rather on the numerical aspects of calculation of the conditional expectation and covariance in the filtering solution derived.

The book at hand is intended for researchers and practitioners in computational and applied science and engineering. However, advanced undergraduate and graduate students with a solid background in linear algebra, vector calculus, Bayesian inference, numerical analysis and with suitable programming skills can also benefit from it. Apart from theoretical developments, this book presents filtering methods under exploration in the form of comprehensive pseudo-codes, which are ready for implementation and use in MATLAB or in any other computation system. Practical performances of state estimators elaborated are demonstrated on stochastic mathematical models in chemistry research, target tracking and electrical engineering.

Strictly speaking, the topic covered in this book is far from new and goes back to the middle of the twentieth century, when it was initiated by the pioneering works of Stratonovich in the 50s–60s and the celebrated paper of Kalman published in 1960. However, what is actually novel and innovative and distinguishes our research from other state estimation studies is regarding numerical integration means as filtering methods. This interpretation allows many achievements in the theory of numerical methods for ordinary differential equations, such as automatic step size selection and error control, to be effectively accommodated to state estimation procedures. It boosts the quality and reliability of filters for estimating continuous-time stochastic systems and gives rise to the new notion of accurate state estimation. Practically, our adaptive filters are devised to be self-turned computational algorithms, which demand no user's effort for discretizing any continuous-time stochastic model at hand with preassigned accuracy. In turn, such methods resolve the problem of sparse and ill-conditioned measurements, automatically. Also, the designed filters admit naturally stochastic systems with missing/irregular sampling. However, the outstanding contribution of this book is state estimation tools for treating stiff continuous–discrete stochastic systems, that is, when the drift coefficient of the stochastic differential equation exposes a stiff behavior. Such stochastic systems constitute a peculiar family of difficult problems, for which traditional filtering means are hardly efficient, because the stiffness can increase the state estimation uncertainty and, hence, its error. On the other hand, stiff models arise often in a number of investigations as, for instance, the Van der Pol oscillator in electrical engineering and the Oregonator reaction in chemistry research. That is why stiff continuous–discrete stochastic systems are of special interest and deserve particular attention in our book.

To summarize, in contrast to other existing studies in filtering theory, this consideration cuts across applied science and engineering (state estimation) and computational mathematics (numerical solution of ordinary differential equations). The basic approach is to employ highly accurate and efficient numerical integration methods available presently for decent predictions of the state expectation and covariance in time-update iterations of the Gaussian filters under exploration. More formally, our focus here is on automatic step size selection and error control facilities underlying the methods in use, which result in self-turned state estimation algorithms where the discretization meshes are not prefixed but, alternatively, generated automatically by the filter itself and with no user's effort required. The latter characteristic of the

techniques under consideration is a necessary prerequisite for an efficient estimation in stiff stochastic systems. All this entails that the reading of the present material demands prior basic knowledge on a broad set of subjects related to numerous issues of numerical integration of ordinary differential equations. Because there exist a severely limited amount of researchers who are equally professional in both the theory of Kalman filtering and that addressing topics on contemporary numerical integration, the first part of our book provides a brief insight into the theory of numerical integration of ordinary differential equations. Structurally, it is split into two chapters. The first one is devoted to the minimal foundations of the aforementioned theory, that is, it gives precise definitions and explains all basic issues and notions of numerical integration in simple terms. These include the concepts of stepping formulas, their consistency, convergence and stability. A special emphasis is paid to implementation aspects of implicit numerical schemes, including the accuracy and stability of the fixed-point iteration as well as those of various Newton-type iterations. The important families of reversible and Hamiltonian problems are also covered, here. In addition, Chap. 1 completes by raising the accuracy of one-step methods via the Richardson extrapolation technique and elaborates all technical details of contemporary local and global error control mechanisms exploited for achieving a desirable accuracy of numerical integration in automatic mode. The presented theoretical analysis of stepping methods is accompanied by pseudo-codes, which clarify the algorithmic specifics of such numerical integration procedures.

Chapter 2 of Part I presents a brief insight into the theory of Runge–Kutta methods for solving ordinary differential equations. More precisely, it gives precise definitions and explains all basic issues and notions in such numerical integration tools. These include the concept of Runge–Kutta formula, its consistency, convergence and stability. The important families of Runge–Kutta schemes intended for integrating reversible and Hamiltonian problems are also covered, here. A special emphasis is paid to the implementation aspects of four subclasses in the realm of implicit Runge–Kutta methods, which allow cheap and efficient Newton-type iterations to be applied. Furthermore, Chap. 2 focuses on variable-stepsize implementations of stepping formulas under exploration and discusses all technical details of contemporary local and global error control mechanisms used for achieving a desirable accuracy of adaptive numerical integration in automatic mode. The outlined analysis of Runge–Kutta formulas is accompanied by pseudo-codes, which clarify the algorithmic specifics of such numerical integration procedures. Finally, an exhaustive theoretical investigation and proper numerical examination of five ODE solvers employed commonly in nonlinear Kalman filtering are fulfilled in MATLAB. To conclude, the first part of our book presents concisely the numerical integration theory at some level suitable for researchers and practitioners working in the stochastic system state estimation area as well as for advanced undergraduate and graduate students learning this subject. Moreover, the readers possessing a solid background in the mentioned theory may skip Part I with no affect to understanding Kalman filtering means themselves, which are addressed in the remaining chapters of our study.

Structurally, Part II of this book is split into five separate chapters devoted to a number of state estimation topics starting off at the most trivial linear discrete-time filtering solution devised by Kalman in 1960 till the more complicated notions of the universal nonlinear continuous–discrete Gaussian filtering algorithms with deterministically sampled expectation and covariance and stiff problems. Chapter 3 presents a brief insight into the theory of the traditional Kalman filtering for discrete-time linear stochastic models. In particular, it gives precise definitions and explains all basic issues and concepts of state estimation in Gauss–Markov stochastic processes. A special emphasis is placed on implementation aspects of the Kalman filtering since it can expose instabilities in solving real-world state estimation tasks because of round-off operations implemented in computer-based simulations. Under some circumstances, such rounding may affect severely the calculation and result in non-symmetric and/or indefinite covariance matrices yielded, which compromise the theoretical rigor of the Kalman filtering and produce poor state estimates. That is why this chapter pays particular attention to the issue of numerical stability and presents a remedy for treating such a covariance-matrix-symmetry-and-positivity-loss in the fashion of square-root filtering methods. Two square-rooting schemes are explored and justified, here. The theoretical analysis of the traditional Kalman filter implementations under consideration is accompanied by their detailed pseudo-codes, which clarify the algorithmic specifics of such state estimation procedures.

Chapter 4 of Part II presents a deep dive into the theory of extended Kalman filtering for continuous–discrete stochastic systems whose process models are of a continuous-time fashion, whereas the measurement ones are discrete-time. In particular, it gives precise definitions and explains all basic issues and notions of state estimation in nonlinear Gaussian systems of such sort. A special emphasis is placed on two extended Kalman filtering design approaches and on their practical implementation aspects since these can expose instabilities in solving real-world state estimation tasks because of the linearization, discretization and rounding operations implemented in computer-based simulations. Under some circumstances, such numerical integration and round-off errors committed may affect severely the calculation and result in non-symmetric and/or indefinite covariance matrices derived, which compromise the theoretical rigor of the extended Kalman filtering technique and produce poor state estimates. That is why this chapter pays particular attention to the issue of numerical stability and presents a remedy for treating such a covariance-matrix-symmetry-and-positivity-loss in the fashion of square-root extended Kalman filtering. Two specific square-rooting schemes grounded on the Cholesky factorization and SVD are explored and justified, here. The theoretical analysis of the extended Kalman filter implementations under study is accompanied by their pseudo-codes, which clarify the algorithmic aspects of such nonlinear state estimation procedures.

Chapter 5 of Part II presents a sound insight into the theory of unscented Kalman filtering for continuous–discrete stochastic systems. In particular, it gives precise definitions of the unscented transform and explores its expectation and covariance approximation properties, which create a solid theoretical background for designing advanced state estimation techniques intended for treating continuous–discrete nonlinear Gaussian systems. A special emphasis is placed on two

unscented Kalman filtering design approaches and on their practical implementation aspects since these can expose instabilities in solving real-world state estimation tasks because of the expectation and covariance approximation, discretization and rounding operations implemented in computer-based simulations. Under some circumstances, such numerical integration and round-off errors committed may affect severely the calculation and result in non-symmetric and/or indefinite covariance matrices computed, which demolish the theoretical rigor of the unscented Kalman filtering and can even fail this state estimation method itself. That is why Chap. 5 pays particular attention to the issue of numerical stability and presents a remedy for treating such a covariance-matrix-symmetry-and-positivity-loss in the fashion of square-root unscented Kalman filtering. Two specific square-rooting techniques grounded on the Cholesky factorization and SVD are explored and justified, here. The theoretical analysis of the unscented Kalman filter implementations under study is accompanied by their pseudo-codes, which clarify the algorithmic aspects of such methods.

Chapter 6 of Part II presents a deep dive into the new topic of universal Kalman filtering with deterministically sampled expectation and covariance for continuous–discrete nonlinear stochastic systems. In particular, it introduces the notion of the general expectation and covariance evaluation principle inspired by the quadrature and cubature rule approximations of the Gaussian-weighted integrals that arise as well as by the unscented transform discussed in Chap. 5. It creates a solid theoretical background for designing universal state estimation tools in the realm of all existing or future Gaussian filtering algorithms with deterministically sampled expectation and covariance. A special emphasis is placed on two Gaussian filtering with deterministically sampled expectation and covariance design approaches and on their practical implementation and utilization aspects since these can expose instabilities in solving real-world state estimation tasks because of the expectation and covariance approximation, discretization and rounding operations involved in computer-based simulations. Under some circumstances, the numerical integration and round-off errors committed may affect severely the calculation and result in non-symmetric and/or indefinite covariance matrices derived, which demolish the theoretical rigor of the universal Kalman filtering with deterministically sampled expectation and covariance and can even fail such state estimation methods. That is why Chap. 6 pays particular attention to the issue of numerical stability and presents a remedy for treating this covariance-matrix-symmetry-and-positivity-loss in the fashion of square-root overall filtering. Two specific square-rooting schemes grounded on the Cholesky factorization and SVD are explored and justified, here. The theoretical analysis of the universal Kalman filters with deterministically sampled expectation and covariance is accompanied by their detailed pseudo-codes, which clarify the algorithmic aspects of such advanced state estimation procedures.

Chapter 7 of Part II presents the first insight into the novel topic of nonlinear Kalman filtering techniques intended for treating stiff continuous–discrete stochastic systems. In particular, it introduces the notion of stiffness in the stochastic differential equation framework and extends it then to continuous–discrete stochastic state estimation tasks under study. Based on the conventional stability analysis of Dahlquist

elaborated in Chap. 1, our consideration focuses on the stiffness features of the specific ordinary differential equations have arisen and been solved within the time-update iterations in the non-square-root and square-root implementations of the traditional extended Kalman filters as well as those of the universal Gaussian ones with deterministically sampled expectation and covariance. The latter are grounded on the Gauss–Hermite quadrature filter, unscented Kalman filter, third- and fifth-degree cubature Kalman filter and derivative-free extended Kalman filter parameterizations. All this creates a solid theoretical background for devising state estimation techniques effective for treating stiff continuous–discrete stochastic systems in practice. The theoretical analysis of A-stable universal Kalman filtering algorithms with deterministically sampled expectation and covariance is accompanied by their detailed pseudo-codes, which clarify the algorithmic aspects of such state estimation tools.

The content of this book covers more than 30 years of research fulfilled at three universities over the world, namely at the Ulyanovsk State University, at the University of the Witwatersrand and at the University of Lisbon. More precisely, the accomplishments of its Part I are an outgrowth of our investigation performed at the first two universities under the partial financial support of the Russian Foundation for Basic Research (Russian Federation) and the National Research Foundation of South Africa, whereas Part II reflects our success achieved while working at the Higher Technical Institute of the University of Lisbon and under the partial financial support of the Foundation for Science and Technology (Portugal). The main purpose of writing the book is exactly to bridge the Kalman filtering theory with that of the efficient and accurate numerical solutions of ordinary differential equations. These research fields have their natural interrelation and share so many common problems and techniques but are not yet effectively collaborating at present.

Lisbon, Portugal
August 2023

Gennady Yu. Kulikov
Maria V. Kulikova

Contents

Acronyms

3D-CKF	Third-degree cubature Kalman filter/third-degree cubature Kalman filtering
5D-CKF	Fifth-degree cubature Kalman filter/fifth-degree cubature Kalman filtering
A 6.24	Algorithm 6.24
A. 6.25	Algorithm 6.25
A. 6.26	Algorithm 6.26
A. 6.27	Algorithm 6.27
A. 6.28	Algorithm 6.28
A. 6.29	Algorithm 6.29
A. 6.30	Algorithm 6.30
ACCF	Auto-cross covariance function/auto-cross covariance functions
ACCM	Auto-cross covariance matrix/auto-cross covariance matrices
ADL-EKF	Accurate-discretized-linearization extended Kalman filter/ accurate-discretized-linearization extended Kalman filtering
ADL-EUKF	Accurate-discretized-linearization extended–unscented Kalman filter/accurate-discretized-linearization extended–unscented Kalman filtering
ADL-KF	Accurate-discretized-linearization Kalman filter/ accurate-discretized-linearization Kalman filtering
ADL-UKF	Accurate-discretized-linearization unscented Kalman filter/ accurate-discretized-linearization unscented Kalman filtering
AGEC	Absolute global error control
AIEC	Absolute iteration error control
ALEC	Absolute local error control
ASRMSE	Accumulated scaled root mean square error/accumulated scaled root mean square errors
BDF	Backward differentiation formula/backward differentiation formulas
BE	Basic event/basic events
BVP	Boundary value problem/boundary value problems

CCF	Cross covariance function/cross covariance functions
CCM	Cross covariance matrix/cross covariance matrices
CCO	Conditional covariance operator/conditional covariance operators
CDF	Conditional distribution function/conditional distribution functions
CEO	Conditional expectation operator/conditional expectation operators
CKF	Cubature Kalman filter/cubature Kalman filtering
CO	Covariance operator/covariance operators
CPDF	Conditional probability density function/conditional probability density functions
CQKF	Cubature–quadrature Kalman filter/cubature–quadrature Kalman filtering
CRV	Conditional random variable/conditional random variables
CSP	Conditional stochastic process/conditional stochastic processes
CSPPDF	Conditional stochastic process probability density function/ conditional stochastic process probability density functions
CSTR	Continuously stirred tank reactor
CWPA	Continuous Wiener process acceleration
DE	Discretization error/discretization errors
DESIRK	Diagonally extended singly implicit Runge–Kutta
DF	Distribution function/distribution functions
DF-EKF	Derivative-free extended Kalman filter/derivative-free extended Kalman filtering
DIRK	Diagonally implicit Runge–Kutta
DL	Discretized linearization
DL-EKF	Discretized-linearization extended Kalman filter/ discretized-linearization extended Kalman filtering
DL-EUKF	Discretized-linearization extended–unscented Kalman filter/ discretized-linearization extended–unscented Kalman filtering
DL-KF	Discretized-linearization Kalman filter/discretized-linearization Kalman filtering
DL-UKF	Discretized-linearization unscented Kalman filter/ discretized-linearization unscented Kalman filtering
DSDE	Deterministic sample differential equation/deterministic sample differential equations
ECG	Electrocardiogram
EEM	Explicit Euler method
EKF	Extended Kalman filter/extended Kalman filtering
EL	Expected loss
EM	Euler–Maruyama
EO	Expectation operator/expectation operators
ERK	Explicit Runge–Kutta
ESDIRK	Explicit-first-stage singly diagonally implicit Runge–Kutta

FNI	Full Newton iteration/full Newton iterations
FPI	Fixed-point iteration/fixed-point iterations
GE	Global error/global errors
GEC	Global error control
GET	Global error tolerance/global error tolerances
GHQF	Gauss–Hermite quadrature filter/Gauss–Hermite quadrature filtering
GLM	General linear method/general linear methods
GNI	General Newton iteration/general Newton iterations
IE	Iteration error/iteration errors
IEM	Implicit Euler method
IET	Iteration error tolerance/iteration error tolerances
IRK	Implicit Runge–Kutta
IT	Itô–Taylor
IVP	Initial value problem/initial value problems
KF	Kalman filter/Kalman filtering
LD	Linearized discretization
LD-EKF	Linearized-discretization extended Kalman filter/ linearized-discretization extended Kalman filtering
LD-KF	Linearized-discretization Kalman filter/linearized-discretization Kalman filtering
LD-UKF	Linearized-discretization unscented Kalman filter/ linearized-discretization unscented Kalman filtering
LE	Local error/local errors
LEC	Local error control
LET	Local error tolerance/local error tolerances
LM	Linear multistep
LTR	Linearized trapezoidal rule
MDE	Moment differential equation/moment differential equations
MIRK	Mono implicit Runge–Kutta
MoL	Method of lines
MPR	Mid-point rule
MVF	Mean value function/mean value functions
NIRK	Nested implicit Runge–Kutta
ODE	Ordinary differential equation/ordinary differential equations
PDE	Partial differential equation/partial differential equations
PDF	Probability density function/probability density functions
RE	Round-off (or rounding) error/round-off (or rounding) errors
RK	Runge–Kutta
RS	Random sequence/random sequences
RT	Richardson transform
RV	Random variable/random variables
SDE	Stochastic differential equation/stochastic differential equations
SDIRK	Singly diagonally implicit Runge–Kutta
SGEC	Scaled global error control

SIEC	Scaled iteration error control
SIRK	Singly implicit Runge–Kutta
SLEC	Scaled local error control
SNI	Simplified Newton iteration/simplified Newton iterations
SP	Stochastic process/stochastic processes
SPDE	Sigma-point differential equation/sigma-point differential equations
SPDF	Stochastic process distribution function/stochastic process distribution functions
SPPDF	Stochastic process probability density function/stochastic process probability density functions
SR	Square root/square roots
SR-ADL-EKF	Square-root accurate-discretized-linearization extended Kalman filter/square-root accurate-discretized-linearization extended Kalman filtering
SR-ADL-EUKF	Square-root accurate-discretized-linearization extended–unscented Kalman filter/square-root accurate-discretized-linearization extended–unscented Kalman filtering
SR-ADL-KF	Square-root accurate-discretized-linearization Kalman filter/ square-root accurate-discretized-linearization Kalman filtering
SR-ADL-UKF	Square-root accurate-discretized-linearization unscented Kalman filter/square-root accurate-discretized-linearization unscented Kalman filtering
SR-DL-EKF	Square-root discretized-linearization extended Kalman filter/ square-root discretized-linearization extended Kalman filtering
SR-DL-EUKF	Square-root discretized-linearization extended–unscented Kalman filter/discretized-linearization extended–unscented Kalman filtering
SR-DL-KF	Square-root discretized-linearization Kalman filter/square-root discretized-linearization Kalman filtering
SR-DL-UKF	Square-root discretized-linearization unscented Kalman filter/ square-root discretized-linearization unscented Kalman filtering
SR-KF	Square-root Kalman filter/square-root Kalman filtering
SR-LD-EKF	Square-root linearized-discretization extended Kalman filter/ square-root linearized-discretization extended Kalman filtering
SR-LD-KF	Square-root linearized-discretization Kalman filter/square-root linearized-discretization Kalman filtering
SR-LD-UKF	Square-root linearized-discretization unscented Kalman filter/ square-root linearized-discretization unscented Kalman filtering
SR-MDE	Square-root moment differential equation/square-root moment differential equations

SVD	Singular value decomposition/singular value decompositions
TR	Trapezoidal rule
UKF	Unscented Kalman filter/unscented Kalman filtering
URE	Unit round-off (or rounding) error
UT	Unscented transform/unscented transforms

Part I
Numerical Solution of Ordinary Differential Equations

Part I is devoted to the theory of numerical integration of ordinary differential equations and organized as follows. Chapter 1 gives examples of problems to be solved numerically and introduces a brief overview of basic issues and concepts of numerical integration in terms of the first and most trivial solution procedures of low order. Its particular interest is paid to the notions playing a crucial role in the improved implementation of Kalman-like filters considered in Part II. The simple fashion of these methods facilitates a better understanding of the fundamental concepts in the theory of numerical solutions of ordinary differential equations under consideration. Then, Chap. 2 presents advanced numerical integration techniques of high order grounded in the contemporary Runge–Kutta formulas together with their most prominent properties and implementation particulars, which strongly influence the performance of such methods in modeling real-world dynamic phenomena. These computational techniques are extensively used in practice and create a fruitful background for the development of novel efficient state estimation algorithms for continuous–discrete stochastic systems, which are the topic of Part II. Theoretical properties of underlying numerical integration tools implemented for the calculation of the mean and covariance in the continuous–discrete Kalman-like filtering context allow many issues in the performance of such methods to be clearly seen and addressed.

Chapter 1
Basic Issues and Concepts of Numerical Integration

Abstract This chapter presents a brief insight into the theory of numerical integration methods for ordinary differential equations. In particular, it gives precise definitions and explains all basic issues and notions of numerical integration. These include the concepts of stepping formulas, their consistency, convergence and stability. A special emphasis is paid to implementation aspects of implicit numerical schemes, including the accuracy and stability of the fixed-point iteration as well as those of various Newton-type iterations. The important families of reversible and Hamiltonian problems are also covered, here. This chapter pays a particular attention to raising the accuracy of one-step methods via the Richardson extrapolation technique and elaborates all technical details of contemporary local and global error control mechanisms used for achieving a desirable accuracy of numerical integration in automatic mode. The presented theoretical analysis of stepping methods under consideration, which are all summarized in the form of pseudo-codes situated in Appendix of this chapter, is always supported with illustrative calculations performed in MATLAB.

1.1 Introduction

Mathematical models enable us to better understand the world by organizing information. They endow us with frameworks for making more accurate predictions, for formulating better strategies and for designing sophisticated policies. Besides, many real-life phenomena evolve in time and allow for their mathematical representation in the form of *ordinary differential equations* (ODE)

$$\mathbf{x}'(t) = \mathbf{f}\big(t, \mathbf{x}(t)\big), \quad t \in [t_0, t_{\rm end}], \tag{1.1}$$

where t is an independent scalar variable (time), the n-dimensional vector $\mathbf{x}(t) = [\, x_1(t)\ x_2(t)\ \dots\ x_n(t)\,]^\top \in \mathbb{R}^n$ describes the state of the plant at time t, the prime stands for differentiation in time and the dynamic behavior of the model is defined by the right-hand side $\mathbf{f}(t, \mathbf{x}(t)) = [\, f_1(t, \mathbf{x}(t))\ f_2(t, \mathbf{x}(t))\ \dots\ f_n(t, \mathbf{x}(t))\,]^\top : \mathbb{D} \subset \mathbb{R}^{n+1} \to \mathbb{R}^n$, which is a sufficiently smooth vector-function on the domain $\mathbb{D}$. The simulation interval $[t_0, t_{\rm end}]$ is assumed to be limited and given in the ODE (1.1).

G. Yu. Kulikov and M. V. Kulikova, *State Estimation for Nonlinear Continuous–Discrete Stochastic Systems*, Studies in Systems, Decision and Control 539,
https://doi.org/10.1007/978-3-031-61371-5_1

Evidently, the formulated mathematical problem is incomplete and admits multiple solutions, which are unacceptable in the deterministic setting considered in this chapter. This inconsistency is resolved and the practical demand for a unique solution is provided by assigning model's initial values as follows:

$$\mathbf{x}(t_0) = \mathbf{x}_0, \tag{1.2}$$

where the vector $\mathbf{x}_0 = [\, x_1(t_0)\ x_2(t_0)\ \ldots\ x_n(t_0)\,]^\top$ stands for the initial state of the plant, which is supposed to be known at time t_0. Eventually, solving the *initial value problem* (IVP) (1.1), (1.2) for the function $\mathbf{x}(t)$ on the time interval of interest supplies us with information on a future evolution of the process under investigation.

The mathematical model (1.1) and (1.2) is a powerful means in studying various aspects of the surrounding world, which is employed extensively in applied science and engineering. For instance, the well-known Van der Pol oscillator obeys the ODE

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ \mu\big(1 - x_1^2(t)\big)x_2(t) - x_1(t) \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} x_1(t_0) \\ x_2(t_0) \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \tag{1.3}$$

where the constant $\mu > 0$ is a fixed real-valued scalar parameter. This dynamic model produces sustained oscillations and is utilized in various circuit explorations of electrical engineering [17, 79–81]. It is also contributed widely to physical, geophysical and biological sciences [16, 32, 71, 75].

The other famous example represents the well-known Oregonator in chemical research and engineering. This model simulates the chemical dynamics of the oscillatory Belousov–Zhabotinsky reaction, as explained, for instance, in [28, 37, 101]. The underlying Oregonator kinetics is an activator/inhibitor system containing both an autocatalytic step and a delayed negative feedback loop. It is composed of five coupled stoichiometries [28]. According to Field and Noyes [31], this reaction can be described mathematically in the fashion of the following ODE:

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{bmatrix} = \begin{bmatrix} 77.27\big[x_2(t) + x_1(t)\big(1 - 8.375 \cdot 10^{-6} x_1(t) - x_2(t)\big)\big] \\ \big[x_3(t) - x_2(t)\big(1 + x_1(t)\big)\big]/77.27 \\ 0.161\big(x_1(t) - x_3(t)\big) \end{bmatrix} \tag{1.4}$$

with the initial values

$$\begin{bmatrix} x_1(t_0) \\ x_2(t_0) \\ x_3(t_0) \end{bmatrix} = \begin{bmatrix} 4.0 \\ 1.1 \\ 4.0 \end{bmatrix}. \tag{1.5}$$

The IVP (1.4) and (1.5) as well as its modifications were exploited extensively in investigations of Hopf bifurcations, Canard explosions, instability, bistability, chaos and many other chemical and physical phenomena [3, 4, 7, 33, 44, 74, 82, 84].

In celestial mechanics, a crucial role is played by the celebrated Kepler problem. This model describes the motion of two bodies, which interact by a central force (the Newtonian gravity) [36]. Mathematically, this is simulated by the ODE

$$\begin{bmatrix} x_1''(t) \\ x_2''(t) \end{bmatrix} = \begin{bmatrix} -x_1(t)\big(x_1^2(t)+x_2^2(t)\big)^{-3/2} \\ -x_2(t)\big(x_1^2(t)+x_2^2(t)\big)^{-3/2} \end{bmatrix} \text{ with } \begin{bmatrix} x_1(t_0) \\ x_2(t_0) \\ x_1'(t_0) \\ x_2'(t_0) \end{bmatrix} = \begin{bmatrix} 1-e \\ 0 \\ 0 \\ \sqrt{1+e}/\sqrt{1-e} \end{bmatrix}, \tag{1.6}$$

where the real-valued constant parameter $e \geq 0$ is the eccentricity of the orbit. The orbit itself can be elliptic, parabolic or hyperbolic depending on the initial total energy [5]. Evidently, the IVP (1.6) is a second-order ODE, but it can be easily transformed to the higher dimensional one of the form (1.1) by introducing artificial variables for the first derivatives, as follows:

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \\ x_4'(t) \end{bmatrix} = \begin{bmatrix} x_3(t) \\ x_4(t) \\ -x_1(t)\big(x_1^2(t)+x_2^2(t)\big)^{-3/2} \\ -x_2(t)\big(x_1^2(t)+x_2^2(t)\big)^{-3/2} \end{bmatrix} \text{ with } \begin{bmatrix} x_1(t_0) \\ x_2(t_0) \\ x_3(t_0) \\ x_4(t_0) \end{bmatrix} = \begin{bmatrix} 1-e \\ 0 \\ 0 \\ \sqrt{1+e}/\sqrt{1-e} \end{bmatrix}. \tag{1.7}$$

The Kepler problem is one of the most fundamental problems in classical mechanics. Its solution represents a closed orbit for every possible set of initial conditions, that is, it returns to the starting point with the same velocity, and can be employed for simulation of a satellite moving around a planet, a planet around its sun or two binary stars around each other and many other celestial phenomena. In addition, it can be applied to modeling the motion of two charged particles and so on.

Another important source of IVP (1.1) and (1.2) is an implementation of the *method of lines* (MoL), which is often applied to solving time-variate *partial differential equations* (PDE) in practice. The MoL implies a semidiscretization of a given PDE, that is, all spatial derivatives are replaced with their discrete approximations (finite differences) whereas time derivatives stay as they are. Eventually, one derives a large-scale ODE of the fashion (1.1), which should be integrated numerically for yielding an approximation to the solution of the original PDE.

To see how the MoL works in applied science and engineering, let us consider a 2D Brusselator model. According [40, p. 151], this PDE has the form

$$\begin{bmatrix} \partial_t u(t,x,y) \\ \partial_t v(t,x,y) \end{bmatrix} = \begin{bmatrix} 1+u^2(t,x,y)v(t,x,y) - 4.4u(t,x,y) + g_1(t,x,y) \\ 3.4u(t,x,y) - u^2(t,x,y)v(t,x,y) + g_2(t,x,y) \end{bmatrix} \tag{1.8}$$

with the right-hand side entries

$$g_1(t,x,y) = \alpha\big(\partial_{xx}^2 u(t,x,y) + \partial_{yy}^2 u(t,x,y)\big) + f(t,x,y), \tag{1.9}$$

$$g_2(t,x,y) = \alpha\big(\partial_{xx}^2 v(t,x,y) + \partial_{yy}^2 v(t,x,y)\big), \tag{1.10}$$

where $\alpha > 0$ is a fixed parameter and the inhomogeneous term is defined as follows:

$$f(t,x,y) = \begin{cases} 5 & \text{if} (x-0.3)^2 + (y-0.6)^2 \leq 0.1^2 \quad \text{and} \quad t \geq 1.1 \\ 0 & \text{else}\,. \end{cases} \tag{1.11}$$

There are three independent variables t (standing for time) and x, y (standing for space) and two unknown functions $u(t, x, y)$ and $v(t, x, y)$ in Eqs. (1.8)–(1.11). The differentiation operators ∂_t and ∂^2_{xx}, ∂^2_{yy} refer to the first- and second-order partial derivatives of the argument function with respect to time and space. The stated PDE is solved on the square domain: $0 \le t \le 6, 0 \le x \le 1, 0 \le y \le 1$. To make the solution that we are looking for unique, the periodic boundary conditions

$$u(t, x, y) = u(t, x + 1, y) = u(t, x, y + 1), \tag{1.12}$$

$$v(t, x, y) = v(t, x + 1, y) = v(t, x, y + 1) \tag{1.13}$$

are set on the boundaries of the space domain at any time t and the initial values

$$u(0, x, y) = 22y(1 - y)^{3/2}, \tag{1.14}$$

$$v(0, x, y) = 27x(1 - x)^{3/2} \tag{1.15}$$

are assigned at the time $t = 0$ and at every space point (x, y) on the solution domain.

The spacial discretization is fulfilled with use of the central differences of order two. For that, we introduce the equidistant meshes in the x and y directions as follows:

$$\{x_i\}_{i=0}^{L_x} = \{x_i = i\,\tau_x,\ i = 0, 1, \ldots, L_x,\ \tau_x = 1/L_x\}, \tag{1.16}$$

$$\{y_j\}_{k=0}^{L_y} = \{y_j = j\,\tau_y,\ i = 0, 1, \ldots, L_y,\ \tau_y = 1/L_y\}, \tag{1.17}$$

where the positive integers L_x and L_y are supplied for providing a desirable accuracy of the numerical integration by the user, that is, these determine the quantities of subdivision nodes utilized for approximating the solution functions $u(t, x, y)$ and $v(t, x, y)$ on the space domain. The equidistant meshes (1.16) and (1.17) allow for the following function and derivative approximations at the mesh nodes in use:

$$u(t, x_i, y_j) \approx u_{ij}(t), \tag{1.18}$$

$$v(t, x_i, y_j) \approx v_{ij}(t), \tag{1.19}$$

$$\partial^2_{xx} u(t, x_i, y_j) \approx \Delta^2_{xx} u_{ij}(t) = \frac{u_{i-1,j}(t) - 2u_{ij}(t) + u_{i+1,j}(t)}{\tau_x^2}, \tag{1.20}$$

$$\partial^2_{yy} u(t, x_i, y_j) \approx \Delta^2_{yy} u_{ij}(t) = \frac{u_{i,j-1}(t) - 2u_{ij}(t) + u_{i,j+1}(t)}{\tau_y^2}, \tag{1.21}$$

$$\partial^2_{xx} v(t, x_i, y_j) \approx \Delta^2_{xx} v_{ij}(t) = \frac{v_{i-1,j}(t) - 2v_{ij}(t) + v_{i+1,j}(t)}{\tau_x^2}, \tag{1.22}$$

$$\partial^2_{yy} v(t, x_i, y_j) \approx \Delta^2_{yy} v_{ij}(t) = \frac{v_{i,j-1}(t) - 2v_{ij}(t) + v_{i,j+1}(t)}{\tau_y^2}. \tag{1.23}$$

Having replaced the solution functions $u(t, x, y)$ and $v(t, x, y)$ and the spacial derivatives with the aforementioned approximations (1.18)–(1.23) in the PDE (1.8)–(1.11) under exploration, we arrive at the large-scale ODE of the fashion

$$\begin{bmatrix} u'_{ij}(t) \\ v'_{ij}(t) \end{bmatrix} = \begin{bmatrix} 1 + u^2_{ij}(t)v_{ij}(t) - 4.4u_{ij}(t) + \alpha\left(\Delta^2_{xx}u_{ij}(t) + \Delta^2_{yy}u_{ij}(t)\right) + f_{ij}(t) \\ 3.4u_{ij}(t) - u^2_{ij}(t)v_{ij}(t) + \alpha\left(\Delta^2_{xx}v_{ij}(t) + \Delta^2_{yy}v_{ij}(t)\right) \end{bmatrix} \tag{1.24}$$

defined at all the nodes of meshes (1.16) and (1.17). In the right-hand side of ODE (1.24), the term $f_{ij}(t)$ implies the inhomogeneous function (1.11) evaluated at the point (t, x_i, y_j). Also, we remark that while evaluating the central differences (1.20)–(1.23) at the space boundaries the subscripts i and/or j can go outside their ranges set in formulas (1.16) and (1.17), respectively. This situation is resolved by means of the periodic boundary conditions (1.12) and (1.13) imposed on the solution. The initial values of ODE (1.24) are computed by formulas (1.14) and (1.15) at the mesh nodes in use as follows:

$$\begin{bmatrix} u'_{ij}(0) \\ v'_{ij}(0) \end{bmatrix} = \begin{bmatrix} 22y_j(1 - y_j)^{3/2} \\ 27x_i(1 - x_i)^{3/2} \end{bmatrix}. \tag{1.25}$$

Thus, we obtain the large-scale IVP (1.24) and (1.25), which is to be integrated on the time interval $[0, 6]$. We stress that this IVP can be high dimensional. For instance, if one takes 49 subdivision steps in each spatial direction (that is, if $L_x = L_y = 49$ in meshes (1.16) and (1.17)) the size of the resulting ODE (1.24) will be 5000.

The IVP (1.1) and (1.2) can also arise in solving scalar *boundary value problems* (BVP) of the conventional fashion

$$x''(t) = f\left(t, x(t), x'(t)\right), \ t \in [t_0, t_{\text{end}}], \ \text{with } x(t_0) = x_0 \text{ and } x(t_{\text{end}}) = x_{end}. \tag{1.26}$$

What distinguishes the BVP (1.26) from the conventional IVP is the fact that the additional information on the solution's behavior is given at different time instants. This issue is resolved by means of the shooting method, which looks for the value of the derivative $x'(t_0)$ at the left boundary t_0 (or, alternatively, for the value of the derivative $x'(t_{\text{end}})$ at the right boundary t_{end}) that ensures the other boundary condition $x(t_{\text{end}}) = x_{end}$ (or $x(t_0) = x_0$). Further details on implementation of the shooting methods for solving the BVP arisen, for instance, in the Cahn–Hilliard continuum modeling of multi-phase fluids can be found in [53] and references therein.

The IVP (1.1) and (1.2) plays a crucial role in filtering theory as well. The latter issue is elaborated in detail and constitutes the topic of Part II in this book.

Eventually, we see that the accurate and efficient solution of the mentioned problem is a question of vital importance in various fields of applied science and engineering because it influences considerably prediction accuracies and, hence, can reduce human power for formulating better strategies and designing sophisticated policies. Furthermore, the above-presented examples of continuous-time dynamic systems exhibit that the IVP (1.1) and (1.2) arisen can be strongly nonlinear and large-scale. This excludes completely any closed-form solution and numerical integration methods should be applied. In the remaining sections of this chapter, we present a brief overview of basic issues and concepts of the theory of numerical methods for ODE in terms of the first and most trivial solution procedures of low order [11, 20, 34, 38–40, 86, 87]. Our particular interest is paid to the notions playing a crucial role

in the improved implementation of Kalman-like filters studied in Part II, below. The simple manner of these methods facilitates a better understanding of the fundamental concepts in the elaborated theory. Then, Chap. 2 presents advanced numerical integration techniques of high-order grounded on contemporary Runge–Kutta formulas together with their most prominent properties and implementation particulars, which strongly affect the performance of such methods, including within stochastic state estimation algorithms considered in the second part of our book.

1.2 Stepping Methods of Low Order

There exist a large variety of different techniques and approaches to integrating the IVP in practice, numerically. It can be solved either in its original differential fashion (1.1) and (1.2) or be translated into the mathematically equivalent integral form

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}\big(s, \mathbf{x}(s)\big)\, ds \tag{1.27}$$

at first. Then, numerical methods suitable for solving this integral equation are applied. In what follows, we focus on stepping methods for integrating the IVP (1.1) and (1.2) because these are more flexible and effective in comparison to quadrature or cubature rules implemented usually for treating the integral problem (1.27).

All stepping schemes are grounded on the idea of replacement of the continuous-time solution function $\mathbf{x}(t)$ defined on the entire integration interval $[t_0, t_{\text{end}}]$ with a sequence of vectors

$$\{\mathbf{x}_l\}_{l=0}^{L} = \{\mathbf{x}_l,\ l = 0, 1, \ldots, L\} \tag{1.28}$$

evaluated at particular time instants

$$\{t_l\}_{k=0}^{L} = \{t_{l+1} = t_l + \tau_l,\ l = 0, 1, \ldots, L-1,\ t_{\text{end}} = t_L\} \tag{1.29}$$

in the aforementioned integration interval. Sequences (1.28) and (1.29) are referred to as a *numerical solution* to the IVP (1.1) and (1.2) and a *numerical integration mesh* assigned in the time interval $[t_0, t_{\text{end}}]$. The *diameter* of mesh (1.29) is its maximum step size, that is,

$$\tau = \max_{l=0,1,\ldots,L-1} \{\tau_l\}. \tag{1.30}$$

Any *stepping method* is a preassigned rule that allows the solution vector $\mathbf{x}_{l+1}$ to be calculated at the subsequent mesh point t_{l+1} by means of the numerical solutions $\mathbf{x}_{l+1-i}$ already computed and available at the corresponding mesh nodes t_{l+1-i}, $i = 1, 2, \ldots, m$, and the size τ_l of the $(l+1)$st step. When $m = 1$ then the stepping formula in use is termed *one-step*. Otherwise, it is referred to as *multistep*. Thus, given the initial value (1.2) and the integration mesh (1.29) each stepping method deter-

mines the unique numerical solution (1.28), which approximates the exact solution $\mathbf{x}(t)$ to the IVP (1.1) and (1.2), when diameter (1.30) of the utilized mesh is sufficiently small. In what follows, we focus on one-step numerical integration procedures.

Clearly, apart from the stepping formula itself, the derived numerical solution $\mathbf{x}_l$, $l = 1, 2, \ldots, L$, depends essentially on the assigned mesh nodes, at which these solution vectors are computed. Such a mesh can be either *variable* (that is, when $\tau_{l+1} \neq \tau_l$ at least for one subscript l) or *equidistant* (that is, when $\tau_{l+1} = \tau_l = \tau$ for all $l = 0, 1, \ldots, L-1$). Any equidistant mesh (1.29) is defined uniquely by its fixed step size $\tau = (t_{\text{end}} - t_0)/L$, which is equal to the diameter of the equidistant mesh, or, this is the same, by the number $L + 1$ of mesh points employed in the integration interval $[t_0, t_{\text{end}}]$. Variable meshes are used basically within *adaptive* stepping schemes. In this case, both the numerical solution (1.28) and its mesh (1.29) are generated by the IVP solver itself in automatic mode, that is, no human effort is required for that. In contrast, equidistant meshes are usually fixed by hand, when no information on the solution's behavior is known in advance and can be taken into account. In the second case, one deals with a *fixed-stepsize* numerical integration procedure. Evidently, the variable-stepsize (or adaptive) numerical integration is more accurate and efficient because it follows the behavior of the exact solution to the IVP (1.1) and (1.2), which we are interested in, and, hence, can regulate its accuracy, automatically. Later on, we address this issue with all the necessary details.

For deriving one-step stepping methods of low order, one can apply the following simple scheme. First, in the concept of stepping procedures, a numerical solution vector $\mathbf{x}_l$ is considered to be known at a current mesh node t_l. The task is to calculate the solution vector $\mathbf{x}_{l+1}$ with the step size $\tau_l = t_{l+1} - t_l$ at the next point t_{l+1} of mesh (1.29). Second, to replace the continuous-time IVP with a discrete-time formula, the derivative $\mathbf{x}'(t)$ in the ODE (1.1) should be approximated at first. In each step of such a method, one has two solution values $\mathbf{x}_l$ and $\mathbf{x}_{l+1}$ available at the mesh nodes t_l and t_{l+1}, respectively. Then, the most evident approach is to exploit the first-order divided difference for yielding the mentioned derivative in every step of mesh (1.29) by the formula

$$\mathbf{x}'(t) \approx (\mathbf{x}_{l+1} - \mathbf{x}_l)/\tau_l, \quad t \in [t_l, t_{l+1}]. \tag{1.31}$$

Third, in line with the ODE (1.1), the derivative approximation (1.31) must equal the ODE's right-hand side evaluated at some point $t^* \in [t_l, t_{l+1}]$. Eventually, choosing different reference points t^* one arrives at the following four stepping formulas:

- Taking $t^* = t_l$ in each step results in the *explicit Euler method* (EEM)

$$\mathbf{x}_{l+1} = \mathbf{x}_l + \tau_l \mathbf{f}\big(t_l, \mathbf{x}_l\big), \quad l = 0, 1, \ldots, L-1. \tag{1.32}$$

- Taking $t^* = t_{l+1}$ in every step leads to the *implicit Euler method* (IEM)

$$\mathbf{x}_{l+1} = \mathbf{x}_l + \tau_l \mathbf{f}\big(t_{l+1}, \mathbf{x}_{l+1}\big), \quad l = 0, 1, \ldots, L-1. \tag{1.33}$$

- With use of $t^* = [t_l + t_{l+1}]/2$, one derives either the *mid-point rule* (MPR)

$$\mathbf{x}_{l+1} = \mathbf{x}_l + \tau_l \, \mathbf{f}\big([t_l + t_{l+1}]/2, [\mathbf{x}_l + \mathbf{x}_{l+1}]/2\big), \quad l = 0, 1, \ldots, L-1, \tag{1.34}$$

or the *trapezoidal rule* (TR)

$$\mathbf{x}_{l+1} = \mathbf{x}_l + \tau_l \left[\mathbf{f}\big(t_l, \mathbf{x}_l\big) + \mathbf{f}\big(t_{l+1}, \mathbf{x}_{l+1}\big)\right]/2, \quad l = 0, 1, \ldots, L-1. \tag{1.35}$$

We stress that the EEM is the oldest numerical integration tool, which has been in use starting at famous Leonhard Euler's studies published in eighteenth century [11].

The one-step methods (1.32)–(1.35) produce unique numerical solution sequences (1.28) for a sufficiently smooth IVP (1.1) and (1.2) on a given mesh of the fashion (1.29) provided that the mesh diameter τ is sufficiently small. Besides, these solutions will be different even if the same mesh (1.29) is utilized in the numerical integrations fulfilled by the above-listed stepping formulas. The latter is because of their different accuracy and other properties of the constructed numerical schemes.

Clearly, the derivative approximation (1.31) entails that the exact solution $\mathbf{x}(t)$ of IVP (1.1) and (1.2) differs from what is calculated by the stepping methods (1.32)–(1.35) in practice. Thus, the crucial property of any numerical solution $\{\mathbf{x}_l\}_{l=0}^{L}$ is its deviation characterized by the following sequence of error vectors:

$$\{\mathbf{ge}_l\}_{l=0}^{L} = \big\{\mathbf{ge}_l = \mathbf{x}(t_l) - \mathbf{x}_l, \; l = 0, 1, \ldots, L \big\} \tag{1.36}$$

evaluated at all nodes of the numerical integration mesh $\{t_l\}_{l=0}^{L}$. We remark that the initial error $\mathbf{ge}_0$ is always zero in this book because the initial values (1.2) are known precisely. Sequence (1.36) is referred to as the *global error* (GE) of the stepping method. It is also known as the *discretization error* in literature. The outcome numerical solution (1.28) is considered to be acceptable when the committed GE is negligible (or reasonably small, that is, it corresponds to the requested accuracy of calculation).

To decide whether the committed GE (1.36) small or large, one has to measure the vector sequence (1.36) in some norm. In theoretical science, one of the most common choices is the use of sup-norm, which results in the error estimate

$$\|\{\mathbf{ge}_l\}_{l=0}^{L}\|_\infty \equiv \max_{l=0,1,\ldots,L} \|\mathbf{ge}_l\|_\infty = \max_{l=0,1,\ldots,L} \max_{i=1,2,\ldots,n} |x_i(t_l) - x_{il}|, \tag{1.37}$$

where the subscript i refers to a particular entry in the corresponding exact and numerical solution vectors and n is the size of the given IVP (1.1) and (1.2). The GE estimate (1.37) is easily interpreted because it shows the maximum deviation of the output numerical solution from the exact one in all entries of the solution vectors.

Unfortunately, the aforementioned *absolute* GE measurement method does not suit for practical implementation. This is due to the maximum magnitude GE entry search, which underlies the error evaluation formula (1.37). In practice, it means that dissimilar physical values can be compared while looking for such an entry. Obvi-