CWI Monographs 2

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

Stability of Runge-Kutta methods for stiff nonlinear differential equations

K. Dekker

J.G. Verwer

North-Holland

CWI Monograph

Stability of Runge-Kutta methods for stiff nonlinear differential equations

K. Dekker

J.G. Verwer

1984

North-Holland
Amsterdam · New York · Oxford

ISBN: 0 444 87634 0

Publishers:
Elsevier Science Publishers B.V.
P.O. Box 1991
1000 BZ Amsterdam
The Netherlands

Sole distributors for the U.S.A. and Canada:
Elsevier Science Publishing Company, Inc.
52 Vanderbilt Avenue
New York, N.Y. 10017
U.S.A.

Cover: © Tobias Baanders 1984

Library of Congress Cataloging in Publication Data

Dekker, K. (Kees)
Stability of Runge-Kutta methods for stiff nonlinear differential equations.

(CWI monograph ; 2)
Bibliography: p.
Includes index.
1. Differential equations, Nonlinear--Numerical solutions. 2. Stiff computation (Differential equations) 3. Runge-Kutta formulas. I. Verwer, J. G. (Jan G.), 1946- . II. Title. III. Series.
QA372.D39 1984 515.3'55 84-18692
ISBN 0-444-87634-0

Printed in the Netherlands

PREFACE

In the last decade an outstanding problem in the field of stiff ordinary differential equations has been to provide an error and stability analysis for nonlinear problems. The developments were initiated by Dahlquist with his lecture at the 1975 Dundee Conference on Numerical Analysis where he introduced the concept of G-stability for multistep methods. Also in 1975, Butcher published his paper on B-stability which deals with Runge-Kutta methods. Both these concepts, which are intimately related, generalize the fundamental property of A-stability to nonlinear systems satisfying a one-sided Lipschitz condition. The introduction of the one-sided Lipschitz condition into the analysis of numerical methods for stiff systems has brought us a long step forward, since for many stiff problems one-sided Lipschitz constants exist which are essentially smaller than the large, positive classical Lipschitz constant. An important theoretical consequence is that we can now derive à priori global error estimates which need not be accompanied by the unrealistic classical assumption that the product of the stepsize and the Lipschitz constant is sufficiently small.

Since Dahlquist and Butcher invented their concepts of G-stability and B-stability there has been much interest in nonlinear stability properties of numerical methods for stiff differential systems. This has led to considerable progress towards a satisfactory theory for numerical methods for stiff nonlinear problems. Without doubt the developments have provided us with much more insight and have led to a much better understanding of the performance of the various methods. The object of the present monograph is to present a unified account of these developments, although the text is confined to the well known class of one-step Runge-Kutta methods. This has to do with our personal interests of course, but also with our aim to present a research monograph of a moderate size.

A first version of this text was presented by one of the authors at a seminar organized in September 1982 by Professor J.M. Sanz-Serna at the University of Bilbao in Spain. The seminar was attended by students and professionals in numerical analysis from the Universities of Bilbao, Valladolid and Zaragoza. This final text is also aimed at students and professionals in numerical analysis. We have tried to make the material accessible to the less experienced reader as well, though it is not meant as an introduction to Runge-Kutta methods. Nor did we attempt to cover the whole theory of Runge-Kutta methods. For example, important topics such at Butcher series, order stars, and codes are not being treated. The text aims at students and scientific researchers who already have some basic knowledge of stiff differential

equations and of Runge-Kutta methods, and of course, a thorough training in applied mathematics is a prerequisite. We hope that the material we make available and our way of presenting it will deepen the reader's understanding of solving numerically stiff differential equations.

The authors wish to emphasize that a book of this kind must owe a great deal to research papers published by others. We therefore acknowledge all those colleagues who have contributed to the subject of this book. In preparing the manuscript the authors have benefited greatly from the willingness of Dr. R. Frank and Professor M. van Veldhuizen for reading parts of an early version. At this place we also mention that part of Chapter 8 emanates from an unpublished note of Professor Van Veldhuizen. We gratefully acknowledge him for supplying this note and for his permission to use it in this text. Special thanks are due to Mrs. M. Louter-Nool for programming the numerical experiments and for checking conscientiously all formulas and proofs. She also assisted in preparing the indexes and in the final proof reading. We are greatly indebted to Dr. L. Shampine who was willing to correct the whole English text. He has provided us with an extensive list of remarks, corrections and improvements. Lastly we acknowledge Mrs. L. Brown and Mrs. J. Kustina for typing the manuscript.

Kees Dekker, Jan G. Verwer

Amsterdam, October 1983.

CWI Monographs

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The CWI is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Stability of Runge-Kutta methods for stiff nonlinear differential equations

CONTENTS

SURVEY

The central theme of this monograph is the step-by-step stability of one-step Runge-Kutta methods for stiff nonlinear ordinary differential systems. The object is to present a unified account of the developments concerning nonlinear stability of Runge-Kutta methods which began with Dahlquist's G-stability paper in 1975. The text is divided into ten chapters.

Chapter 1 is devoted to the stiff problem itself. There we define the initial value problem we are dealing with and recall the phenomenon of stiffness. Further we introduce the one-sided Lipschitz constant and the more general concept of logarithmic matrix norm. Nowadays both these quantities are considered to be indispensable mathematical tools in the perturbation analysis of stiff equations, analytically as well as numerically. Loosely speaking, the logarithmic norm, which is just an optimal one-sided Lipschitz constant when dealing with inner product norms, replaces the classical Lipschitz constant. For stiff problems the latter constant is too large to be of any use in a perturbation analysis. It must be assumed here, of course, that the logarithmic matrix norm can be chosen essentially smaller than the classical Lipschitz constant. The logarithmic matrix norm was introduced independently by Dahlquist in 1958 (see Dahlquist [1959]) and Lozinskij [1958]. Since then it has been used in various studies, most of which deal with ordinary differential equations. Our treatment of the one-sided Lipschitz constant and the logarithmic matrix norm is not exhaustive, but it is certainly sufficient for a first study. Much of the material of Chapter 1 is based on Dahlquist [1959], Desoer and Haneda [1972], Frank, Schneid and Ueberhuber [1978], and Ström [1975].

In Chapter 2 we discuss the concepts of numerical contractivity and numerical stability from a general point of view. Numerical contractivity means that the difference, in a certain norm, of any two numerical solutions defined for equal stepsizes will not increase as the integration proceeds. With numerical stability we are referring to a weaker and less precisely defined numerical property. Numerical stability means that in a sense the aforementioned difference may grow as the integration proceeds, but not unboundedly. Here we distinguish convergence stability and computing stability which is, in fact, contractivity. By way of introduction we revert to the classical concept of absolute stability and review some important results for the constant coefficient linear model system emanating from Von Neumann [1951], Hairer, Bader and Lubich [1982] and Spijker [1983a]. Next, a rather exceptional contractivity property of the well known implicit Euler method which goes back to Desoer and Haneda [1972] is proved. For implicit Euler we then illustrate the advantage of using the logarithmic matrix norm instead of the classical Lipschitz constant. The

classical Lipschitz constant is large when the differential equation is stiff, hence of no use for the numerical analysis. We also derive a general à priori error bound for an arbitrary one-step integration method from which the importance of contractivity and nonlinear stability can be explained. Chapter 2 concludes with an instructive example illustrating the fact that as soon as one leaves the class of constant coefficient systems, one cannot rely on the spectrum of the Jacobian matrix to infer reliable information about the error propagation (cf. Lambert [1978]).

In Chapter 3 we review the numerical methods we concentrate on, viz. the one-step Runge-Kutta methods. All the properties of the Runge-Kutta methods needed in subsequent chapters have been brought together in this chapter. The emphasis lies on quadrature type methods. However we also pay attention to singly-implicit and diagonally implicit methods. These two latter kinds of methods are normally regarded as more attractive for computer implementation than the standard quadrature methods such as Gauss-Legendre, Radau and Lobatto. It will be clear that many parts of Chapter 3 emanate from the pioneering work of Butcher (see the Bibliography). One section is entirely devoted to the W-transformation of Hairer and Wanner [1981]. This transformation greatly simplifies the classification and analysis of Runge-Kutta formulas.

Chapter 4 is completely devoted to the closely related topics of B-stability, BN-stability, AN-stability and algebraic stability. Consequently, in this chapter we often quote from Butcher [1975], Burrage and Butcher [1979], and Crouzeix [1979]. We remark that B-stability, BN-stability and AN-stability are in fact contractivity properties. A great deal of attention is paid to the relation between these topics which means that we also describe the concepts of reducibility as defined by Dahlquist and Jeltsch [1979] and Hundsdorfer and Spijker [1981a].

Chapter 5 deals with the existence and uniqueness of solutions of the systems of nonlinear algebraic or transcendental equations which arise in the application of the implicit Runge-Kutta methods. In most of the literature on nonlinear stability and contractivity, one assumes in advance that existence and uniqueness of solutions to these systems is ensured. This assumption may be false, however; Hundsdorfer and Spijker [1981b] give an example which shows that algebraic stability and dissipativity are not sufficient. In general, additional conditions must be imposed to ensure existence and uniqueness for dissipative as well as non-dissipative problems. We give a lot of attention to the derivation of these extra conditions. Here we mainly base our treatment on Hunsdorfer and Spijker [1981b], Dekker [1982], and Crouzeix, Hundsdorfer and Spijker [1983]. Furthermore, we also quote from Frank, Schneid and Ueberhuber [1982a], as it has turned out that various results on existence and solvability also play an important role in their work on B-convergence which

we treat in Chapter 7. In the last section of Chapter 5 we briefly discuss aspects of the implementation of implicit Runge-Kutta formulas. An experiment with the singly-implicit code STRIDE developed by Burrage, Butcher and Chipman [1980] is reported.

In Chapter 6 we discuss the concept of circle contractivity introduced by Dahlquist and Jeltsch [1979]. Circle contractivity, or generalized disk contractivity, bears a close resemblance to algebraic stability. Circle contractivity, however, can also be used for studying nonlinear contractivity properties of explicit methods for certain classes of nonstiff problems. We focus our attention on this particular feature. We emphasize that Chapter 6 is completely based on the 1979 report of Dahlquist and Jeltsch.

A well known phenomenon in the numerical integration of stiff differential equations is the often disappointing accuracy of the numerical result when contrasted with the order of consistency of the method. For Runge-Kutta methods Prothero and Robinson [1974] have studied this phenomenon for a simple, but very instructive scalar test problem. To deal with the general case Frank, Schneid and Ueberhuber [1981a] invented the notion of B-convergence which we discuss in Chapter 7. The idea of B-convergence is to derive à priori error bounds which depend solely on the smoothness of the exact solution and on a one-sided Lipschitz constant. Hence, under certain conditions, these bounds are completely independent of the stiffness of the problem. The B-convergence theory gives a great deal of insight into the process of integrating a stiff problem. In particular, B-convergence theory shows that the celebrated properties of superconvergence of quadrature methods is of only limited value for stiff problems. To illustrate this point the chapter presents the results of some numerical experiments. The greater part of Chapter 7 is based on the work of Frank, Schneid and Ueberhuber [1981a,b; 1982a,b]. We note that the subject of Chapter 7 is accuracy rather than stability. The B-convergence theory, however, heavily leans upon various results on nonlinear stability, which makes this B-convergence chapter fit nicely into our monograph.

In Chapter 8 we discuss the D-stability concept which was introduced by Van Veldhuizen [1981]. D-stability is related to a particular class of linear problems containing a small parameter, the stiffness parameter. It is not a stability property in the usual sense of stepwise stability, i.e., it is not related to the evolution in time of the solution of the discretized problem, in contrast to other concepts such as A-stability and B-stability. It is merely a boundedness property of the numerical solution in the presence of a small parameter. In this respect, D-stability is somewhat outside the scope of this monograph. However, D-stability is useful in its own right because it assumes nothing about the existence of a logarithmic norm of moderate size. Proofs of results on D-stability are normally based on asymptotic arguments. The main purpose of this chapter is to explain the concept and to provide some background

material.

Chapter 9 is entirely devoted to Runge-Kutta-Rosenbrock methods. These methods are closely related to diagonally implicit Runge-Kutta methods. Most practical implementations of the latter belong to the class of Runge-Kutta-Rosenbrock methods. Rosenbrock methods do not possess contractivity and nonlinear stability properties as strong as those of most implicit Runge-Kutta methods. Following Hairer, Bader and Lubich [1982], we shall show that under certain conditions these methods are able to produce stable solutions for nonlinear problems when the right hand side function can be split into a stiff linear part and a nonstiff nonlinear part. If such a splitting does not exist, stability cannot be guaranteed. We shall illustrate this by discussing some results on D-stability from Van Veldhuizen [1981] and Verwer [1982a,b]. In this chapter it is also pointed out, following Verwer [1981], that nonlinear stability properties of implicit methods, such as BN-stability, are difficult to exploit when implementing the implicit method on the computer. The source of the difficulty is the iterative numerical solution of the implicitly defined Runge-Kutta approximation. Loosely speaking, the always disappointing difference between theory and practice manifests itself here. The experiment with the singly-implicit code STRIDE presented in Section 5.12 illustrates this point clearly. In Section 9.7 we discuss a similar experiment with the Rosenbrock code ROW4A developed by Kaps and Rentrop [1979] and Gottwald and Wanner [1981].

The last chapter of the monograph, Chapter 10, deals with step-by-step stability in the numerical solution of evolutionary problems for partial differential equations. For many discretizations of such problems the time-integration part appears to be a numerical integration formula for stiff ordinary differential equations. This implies that there must be close connections between stiff problems and partial differential equations with respect to numerical step-by-step stability. Following Verwer and Dekker [1983a], the object of the last chapter is to illustrate these connections. The well known energy method from partial differential equations plays an important role here. At the cost of some repetition the material of Chapter 10 is more or less self-contained.

CHAPTER 1

STIFF DIFFERENTIAL EQUATIONS

This book deals with the numerical solution of the initial value problem for stiff systems of ordinary differential equations. Throughout we shall use

$$\dot{y}(t) = f(t,y(t)), \quad 0 \leqslant t \leqslant T, y(0) = y_0, \tag{1.1}$$

to denote some problem or class of problems under consideration. Here $y(t)$ is a real vector of m elements and f a real-valued vector function, possibly nonlinear in the dependent and independent variables.

For the sake of analysis we introduce an auxiliary function $h: [0,T] \to \mathbb{R}^m$ and assume that any solution y to be considered satisfies

$$\|y(t) - h(t)\| \leqslant \psi(t), \quad 0 \leqslant t \leqslant T, \tag{1.2}$$

where $\psi: [0,T] \to \mathbb{R}$ is appropriately chosen and $\|\cdot\|$ denotes some vector norm on $\mathbb{R}^m$. Associated with the functions h and ψ is the convex region

$$M_t = \{\zeta \in \mathbb{R}^m : \|\zeta - h(t)\| \leqslant \psi(t)\}. \tag{1.3}$$

By definition any solution value $y(t) \in M_t$. The function h is allowed to be a solution of (1.1) itself. In what follows M_t is the domain of the function $f(t,\cdot): M_t \to \mathbb{R}^m$.

It will always be tacitly assumed that f is as often differentiable as the numerical analysis requires. This implies that for all initial vectors $y_0 \in M_0$, problem (1.1) possesses a unique solution for all $t \in [0,T]$. In such a situation it is supposed that T can be chosen as large as we wish, e.g., infinite.

The first chapter is devoted to the stiff problem itself. By way of introduction we shall describe the phenomenon of stiffness in Section 1.1. The remaining sections deal with the one-sided Lipschitz constant and, more generally, with the logarithmic matrix norm. These concepts play an important role in modern numerical literature on stiff problems.

1.1. Stiffness

The problems called stiff are diverse and it is rather cumbersome to give a mathematically rigorous definition of stiffness. Consequently, in the literature various definitions are seen, one being somewhat more precise than another. *The essence of stiffness is that the solution to be computed is slowly varying but that perturbations exist which are rapidly damped.* The presence of such perturbations complicates the numerical computation of the slowly varying solution.

We shall discuss a number of simple examples of stiff problems in order to provide some background for the statements made above.

Example 1.1.1. Following Shampine and Gear [1979], we first consider the scalar equation

$$\dot{y}(t) = \lambda y(t) + \dot{F}(t) - \lambda F(t), \quad t \geq 0, \; y(0) = y_0, \; \lambda \ll 0, \tag{1.1.1}$$

where F is a slowly varying function of t only. The solution $y(t)$ is given by

$$y(t) = F(t) + e^{\lambda t}[y_0 - F(0)],$$

or, relating the solution at time t to time $t+\tau$,

$$y(t+\tau) = F(t+\tau) + e^{\lambda \tau}[y(t) - F(t)]. \tag{1.1.2}$$

Because $\lambda \ll 0$, it is clear that already after a very short time distance the *transient,* $e^{\lambda t}[y_0 - F(0)]$, which is also called the *stiff solution component* or *strongly varying solution component,* is no longer present in the solution $y(t)$. This means that, irrespective of the initial value y_0, the slowly varying function $F(t)$ dominates the solution to be computed on the larger part of the integration interval $[0,T]$. The second expression for the solution shows that at any time t perturbations to the slowly varying solution $F(t)$ exist which are rapidly damped. In passing we note that *slowly varying solution components* are often referred to as *nontransient components* or *smooth components.* The word smooth is used in the sense of the derivatives being essentially smaller than the derivatives of the transient components. □

To expose some basic aspects of solving numerically a super stable problem like (1.1.1), we now apply the well known *Euler* methods

$$y_{n+1} = y_n + \tau f(t_n, y_n), \tag{1.1.3}$$

and

$$y_{n+1} = y_n + \tau f(t_{n+1}, y_{n+1}). \tag{1.1.4}$$

Formula (1.1.3) is the simplest example of a classical *explicit* Runge-Kutta formula (see Henrici [1962]), while (1.1.4) is the simplest *implicit* formula of the Runge-Kutta family. Let us compare the approximations defined by (1.1.3) and (1.1.4) with the exact solution $y(t_{n+1})$, $t_{n+1} = t_n + \tau$, given by (1.1.2). We have

$$y(t_{n+1}) = e^{\lambda \tau}[y(t_n) - F(t_n)] + F(t_{n+1}), \tag{1.1.2}$$

$$y_{n+1} = (1+\tau\lambda)[y_n - F(t_n)] + F(t_n) + \tau \dot{F}(t_n), \tag{1.1.3}$$

$$y_{n+1} = (1-\tau\lambda)^{-1}[y_n - F(t_n)] + \tag{1.1.4}$$

$$(1-\tau\lambda)^{-1}[F(t_n)+\tau\dot{F}(t_{n+1})-\tau\lambda F(t_{n+1})].$$

As explained above, $y(t_n)-F(t_n)$ may be interpreted as a perturbation to the smooth solution $F(t)$ at $t=t_n$. This perturbation is rapidly damped by the negative exponential $e^{\lambda\tau}$, $\lambda\tau\ll 0$. One may interpret the differences $y_n-F(t_n)$ in a similar way. The numerical method should be able to damp the differences for values $\tau\lambda\ll 0$, and τ should be adjusted only in response to variations in the function $F(t)$.

From (1.1.4′) it is readily seen that implicit Euler performs very satisfactorily on equation (1.1.1). The differences $y_n-F(t_n)$ are rapidly damped and we even find that for any fixed $\tau>0$, $y_{n+1}\to F(t_{n+1})$ as $\tau\lambda\to-\infty$. The explicit Euler method is not capable in solving (1.1.1) efficiently. The difference $y_n-F(t_n)$ is damped only if $-2<\tau\lambda<0$. This condition of numerical stability imposes a severe restriction on the stepsize τ if $\lambda\ll 0$, even when $y_n-F(t_n)$ is negligibly small. On the other hand, if $y_n\simeq F(t_n)$ and F is very smooth, the approximation of $F(t_{n+1})$ by $F(t_n)+\tau\dot{F}(t_n)$ will be acceptable for much larger values of τ than those admitted by the stability condition $\tau<-2/\lambda$. This situation is typical when applying an explicit method to a stiff problem. *The stepsize is restricted by numerical stability rather than by accuracy.* The super stability of the differential equation turns out to be disadvantageous for the error propagation of classical explicit methods.

Example 1.1.2. The general solution of (1.1.1) contains a strongly varying component, the stiff or transient component, and a weakly varying component, the smooth component. For linear systems $\dot{y}(t)=Ay(t)$, A a constant matrix, such different solution components occur when the Jacobian matrix A possesses eigenvalues which differ greatly in magnitude. Let us consider the linear system

$$\dot{y}(t)=\begin{pmatrix} d & \epsilon^{-1} \\ 0 & -\epsilon^{-1}\end{pmatrix}y(t),\quad t\geqslant 0,\quad y(0)=y_0,\quad \epsilon>0, \tag{1.1.5}$$

where d is a constant of moderate size, $|d|<1$ say, and ϵ is close to zero. We have

$$y(t)=\begin{pmatrix} e^{dt} & (e^{dt}-e^{-\epsilon^{-1}t})/(1+d\epsilon) \\ 0 & e^{-\epsilon^{-1}t}\end{pmatrix}y_0,\quad t\geqslant 0. \tag{1.1.6}$$

The transient component is given by $e^{-\epsilon^{-1}t}$. Since ϵ is very small, this component dies out after a very short time. After this transient phase, the solution is completely determined by the smooth component e^{dt}. In this phase of the problem, we are faced with the task of computing the nontransient solution while suppressing the transient one. The problem is then called *stiff*.

It is emphasized that the problem is not called stiff in the transient phase