P.J. van der Houwen

CONFERENTIE VAN NUMERIEK WISKUNDIGEN

23 - 25 september 1998

CONFERENTIECENTRUM WOUDSCHOTEN
ZEIST

Werkgemeenschap Numerieke Wiskunde

CONFERENTIE
VAN NUMERIEK WISKUNDIGEN

23 - 25 september 1998

CONFERENTIECENTRUM WOUDSCHOTEN
ZEIST

Werkgemeenschap Numerieke Wiskunde

DRIEENTWINTIGSTE CONFERENTIE NUMERIEKE WISKUNDE

Doel van de conferentie

De Conferentie van Numeriek Wiskundigen wordt eenmaal per jaar gehouden onder auspiciën van de Werkgemeenschap Numerieke Wiskunde en de Wetenschappelijke Onderzoeksgemeenschap (WOG) 'Numerieke methoden voor wiskundige modellering' (FWO-Vlaanderen). Het doel van de conferentie is kennis te nemen van recente ontwikkelingen binnen de numerieke wiskunde. Hiertoe worden jaarlijks twee thema's vastgesteld. Internationaal bekende deskundigen worden uitgenodigd over deze thema's lezingen te houden.

Thema's

1. aspecten van de integratie van beginwaardeproblemen
2. wavelets en hiërarchische bases

Organisatie

De organisatie is in handen van de voorbereidingscommissie bestaande uit
P. Wesseling (TUD) (voorzitter), D. Roose (KU Leuven), M. Spijker (RU Leiden), en J. Kok (CWI) (secretaris). Organisatorische medewerking is verleend door het Centrum voor Wiskunde en Informatica. Financiële ondersteuning is gegeven door de Stichting Wiskunde Onderzoek Nederland (SWON) en de WOG.

Adres organisatie

Jan Kok
CWI (cl. MAS)
Postbus 94079
1090 GB Amsterdam
Fax: (020) 592 4199 / +31 20 592 4199
Tel : (020) 592 4107 / +31 20 592 4107
E-mail: `Jan.Kok@cwi.nl`

Conferentie-URL: `http://www.cwi.nl/~jankok/woudschoten.html`

Arrangements

Coffee and tea during breaks will be served in the *foyer*, i.e. the ground level of the central part of the new wing

Coffee at 19.30 after the dinners, and also the first day before the start of the conference: in the *lounge*, i.e. left of the *lobby (*seen from the *entrance*)

Bar in the lounge: open till 21.00,
from 21.00 - 0.30 the bar in the *souterrain* is open (the souterrain is below the foyer)

Meals in the *restaurant*, i.e. right of the *lobby (*seen from the *entrance*)

All guests are requested to vacate their room and return the key on the day of departure BEFORE the start of the first conference session

The conference organization does not take up costs made for the use of telephone, fax, copying machine, and for drinks during dinner or at the bars. All participants are requested to pay for these costs directly with the conference centre staff.

Bus info (bus 81) *(info for 1998)*

From Utrecht (streekbusstation) to Zeist (busstop Woudschoten / Oud London)
10.09, 11.09 (takes 34 minutes)

From Woudschoten to Utrecht (streekbusstation)
15.47, 16.19, 16.47 (takes 35 minutes)

Twenty-third Dutch Conference on Numerical Analysis

Themes and Speakers

Theme 1.	*aspects of the integration of initial-value problems*
	Kevin Burrage (University of Queensland) Andrew Stuart (Stanford University) Marino Zennaro (Università di Trieste)
Theme 2.	*wavelets and hierarchical bases*
	Jean-Pierre Antoine (Université Catholique de Louvain), Wolfgang Dahmen (RWTH Aachen), Peter Oswald (Bell Labs)

Contributed presentations by:

Jos van Dorsselaer, CWI (Theme 1)
Tanja Van Hecke, University of Gent (Theme 1)
Karel in 't Hout, Leiden University (Theme 1)
Rob Stevenson, University of Nijmegen (Theme 2)

Twenty-third Dutch Conference on Numerical Analysis

Programme and titles of lectures

Wednesday, September 23, 1998

10.00 - 11.05	**arrival, coffee**	
11.10	**opening**	
11.15	**J-P. Antoine**	
	11.15 - 12.05:	The continuous wavelet transform, wavelet packets and fast algorithms
	12.05 - 12.15:	discussion
12.20	**lunch**	
13.45	**M. Zennaro**	
	13.45 - 14.35:	An introduction to the numerical solution of delay differential equations
	14.35 - 14.45:	discussion
14.50	**K. Burrage**	
	14.50 - 15.40:	An introduction to the numerical solution of stochastic ordinary differential equations
	15.40 - 15.50:	discussion
15.50	**tea**	
16.15	**W. Dahmen**	
	16.15 - 17.05:	Wavelets and adaptivity in numerical analysis, I
	17.05 - 17.15:	discussion
17.20	**J. van Dorsselaer**	
	17.20 - 17.45:	Inertial manifolds of parabolic pde's under time discretization
	17.45 - 17.50:	discussion
18.15	**dinner**	
20.00	**WNW Committee meeting,** followed by **Woudschoten Committee meeting**	

Thursday, September 24, 1998

08.00	**breakfast**	
09.00	**A. Stuart**	
	09.00 - 09.50:	Long-term integration of stochastic differential equations
	09.50 - 10.00:	discussion
10.00	**coffee**	
10.35	**P. Oswald**	
	10.35 - 11.25:	Multilevel frames and subspace splittings with applications to iterative methods
	11.25 - 11.35:	discussion
11.40	**J-P. Antoine**	
	11.40 - 12.30:	Directional 2-D wavelets and applications
	12.30 - 12.40:	discussion
12.45	**lunch**	
14.25	**M. Zennaro**	
	14.25 - 15.15:	Some stability problems for delay differential equation solvers
	15.15 - 15.25:	discussion
15.30	**T. Van Hecke**	
	15.30 - 15.55:	Deferred Correction with mono-implicit Runge-Kutta methods for first order IVPs
	15.55 - 16.00:	discussion
16.00	**tea**	
16.25	**R. Stevenson**	
	16.25 - 17.15:	Element-by-element construction of wavelets satisfying stability and moment conditions
	17.15 - 17.25:	discussion
17.30	**General Assembly of the Dutch** "Werkgemeenschap Numerieke Wiskunde"	
18.15	**dinner**	

Friday, September 25, 1998

08.00	**breakfast**	
09.00	**K. Burrage**	
	09.00 - 09.50:	Implementation issues in solving stochastic ordinary differential equations
	09.50 - 10.00:	discussion
10.05	**K. in 't Hout**	
	10.05 - 10.30:	The convergence of Runge–Kutta methods for delay differential equations
	10.30 - 10.35:	discussion
10.35	**coffee**	
11.10	**W. Dahmen**	
	11.10 - 12.00:	Wavelets and adaptivity in numerical analysis, II
	12.00 - 12.10:	discussion
12.15	**lunch**	
13.30	**A. Stuart**	
	13.30 - 14.20:	Long-term integration of large coupled systems of oscillators
	14.20 - 14.30:	discussion
14.35	**P. Oswald**	
	14.35 - 15.25:	Multilevel discretization schemes for the single layer potential equation
	15.25 - 15.35:	discussion
15.35	**closure, tea, departure**	

Page numbers of abstracts

	abstract	*theme*	*page*
J-P. Antoine	The continuous wavelet transform, theory and applications	2	8
K. Burrage	An introduction to the numerical solution of stochastic ordinary differential equations	1	12
K. Burrage	Implementation issues in solving stochastic ordinary differential equations	1	13
Wolfgang Dahmen	Wavelets and adaptivity in numerical analysis I, II	2	14
Jos van Dorsselaer	Inertial manifolds of parabolic pde's under time discretization	1	16
T. Van Hecke	Deferred Correction with mono-implicit Runge-Kutta methods for first order IVPs	1	17
Karel in 't Hout	The convergence of Runge–Kutta methods for delay differential equations	1	24
Peter Oswald	Multilevel frames and subspace splittings with applications to iterative methods	2	26
Peter Oswald	Multilevel discretization schemes for the single layer potential equation	2	32
Rob Stevenson	Element-by-element construction of wavelets satisfying stability and moment conditions	2	36
Andrew Stuart	Stiff oscillatory systems with random initial data	1	38
Andrew Stuart	Perturbation theory for ergodic Markov chains	1	44
M. Zennaro	An introduction to the numerical solution of delay differential equations	1	48
M. Zennaro	Some stability problems for delay differential equation solvers	1	52

The Continuous Wavelet Transform, theory and applications

J-P. Antoine
Institut de Physique Théorique
Université Catholique de Louvain
B–1348 Louvain-la-Neuve, Belgium
e-mail: antoine@fyma.ucl.ac.be

Lecture #1 – The CWT : definitions, implementation, applications

The one-dimensional wavelet transform has found nowadays many applications to various fields of physics, mathematics and signal processing. The original motivation was to design a method of analysis suitable for nonstationary, highly inhomogeneous signals (such as speech), for which Fourier analysis is inadequate. The outcome is a *time-scale* analysis, based on the wavelet transform (WT):

$$S(b,a) = a^{-1/2} \int \overline{\psi\left(a^{-1}(t-b)\right)}\, s(t)\, dt \equiv \langle \psi_{b,a} | s \rangle, \tag{1}$$

where $a > 0$ is a scale parameter and $b \in \mathbb{R}$ a translation parameter. In the relation (1), s is a finite energy signal, the function ψ, the analyzing wavelet, is assumed to be well localized *both* in the time domain and in the frequency domain, and the bracket denotes the usual scalar product in $L^2(\mathbb{R}, dt)$. In addition ψ must satisfy an admissibility condition, which in most cases may be reduced to the requirement that ψ has zero mean (hence it is sufficiently oscillating): $\int \psi(t)\, dt = 0$. Combining this condition with the localization properties of $\psi(t)$ and its Fourier transform $\widehat{\psi}(\omega)$, one sees that the WT $s \mapsto S$ provides a *local filtering*, both in time (b) and in scale (a), which works at constant relative bandwidth, $\Delta\omega/\omega =$ constant. Thus it is more efficient at high frequency, i.e. small scales, in particular for the detection of singularities in the signal. In addition, the transformation $s(x) \mapsto S(a,b)$ may be inverted exactly and yields a reconstruction formula, which amounts to a decomposition of the signal in terms of dilated, translated copies $\psi_{b,a}$ of the basic wavelet ψ.

Of course, the numerical implementation requires the discretization of integrals. In particular, the reconstruction formula expresses the signal as a linear superposition of a discrete family $\{\psi_{b_i,a_j}\}$. However, in general, this approach does *not* lead

to an orthonormal basis. In order to achieve this, it is necessary to exploit a totally different approach, based on *multiresolution analysis*, thus leading to the discrete wavelet transform (DWT).

Both the DWT and the continuous wavelet transform (CWT) extend to 2 (or more) dimensions, with exactly the same properties as in the 1-D case. Here again the mechanism of the WT is easily understood from its very definition as a convolution:

$$S(\vec{b}, a, \theta) = a^{-1} \int \overline{\psi\left(a^{-1} r_{-\theta}\left(\vec{x} - \vec{b}\right)\right)} s(\vec{x})\, d^2\vec{x}, \tag{2}$$

where s is the signal and ψ is the analyzing wavelet, which is translated by $\vec{b} \in \mathbb{R}^2$, dilated by $a > 0$ and rotated by an angle θ ($r_{-\theta}$ is the rotation operator). Since the wavelet ψ is required to have zero mean, we have again a filtering effect, i.e. the analysis is *local* in all four parameters $\vec{b}, a, \theta$, and here too it is particularly efficient at detecting discontinuities in images. When compared to the 1-D case, the new fact here is the presence of the rotation degree of freedom. This is crucial for detecting *oriented features* of the signal, that is, regions where the amplitude is regular along one direction and has a sharp variation along the perpendicular direction, for instance, edges or contours . The CWT is a very efficient tool in this respect, provided one uses a *directional* wavelet, that is, a wavelet which has itself an intrinsic orientation (for instance, it contains a plane wave).

It is a quite common opinion that the CWT is too time consuming for any practical use in image processing. This is, we think, a misconception. Not only is it better adapted in a number of situations, but in addition fast algorithms have been designed recently that make it truly competitive numerically.

In this first lecture, we will survey the theory and some applications of the continuous WT, both in 1-D and in 2-D. The following points will be developed:

1. Definition and general properties: covariance, norm conservation, reconstruction formula (inverse CWT), reproducing kernel.

2. Interpretation of the CWT as a singularity scanner.

3. Practical implementation of the 2-D CWT: The position and scale-angle representations; standard wavelets.

4. Discretization of the CWT, comparison with the DWT.

5. Fast algorithms for the CWT: continuous wavelet packets, pseudo-QMFs.

6. Some applications of the CWT, mostly in 2-D.

Lecture #2 – Directional 2-D wavelets and applications

As mentioned above, the analysis of oriented features in an image requires the use of the full 2-D CWT, including the rotation degree of freedom, and also a *directional* wavelet ψ. By this we mean that the effective support of its Fourier transform $\widehat{\psi}$ is contained in a convex cone in spatial frequency space $\{\vec{k}\}$, with apex at the origin. Two standard examples are :

(i) *The Morlet wavelet:*

$$\psi_M(\vec{x}) = \exp(i\vec{k}_o \cdot \vec{x})\, \exp(-\tfrac{1}{2}|A\vec{x}|^2), \quad \widehat{\psi}_M(\vec{k}) = \sqrt{\epsilon}\, \exp(-\tfrac{1}{2}|A^{-1}(\vec{k} - \vec{k}_o)|^2), \quad (3)$$

The parameter $\vec{k}_o$ is the wave vector, and $A = \mathrm{diag}[\epsilon^{-1/2}, 1], \epsilon \geq 1$, is a 2×2 anisotropy matrix.

(ii) *The Cauchy wavelet:*

$$\widehat{\psi}^{(C)}_{lm}(\vec{k}) = \begin{cases} (\vec{k} \cdot \vec{e}_{\tilde{\alpha}})^l\, (\vec{k} \cdot \vec{e}_{-\tilde{\alpha}})^m\, e^{-\vec{k}\cdot\vec{\eta}}, & \vec{k} \in \mathcal{C}(-\alpha, \alpha) \\ 0, \quad \text{otherwise.} \end{cases} \quad (4)$$

Here $\mathcal{C} \equiv \mathcal{C}(-\alpha, \alpha) = \{\vec{k} \in \mathbb{R}^2 \mid -\alpha \leq \arg \vec{k} \leq \alpha\}$ is the convex cone determined by the unit vectors $\vec{e}_{-\alpha}, \vec{e}_{\alpha}$, $\tilde{\alpha} = -\alpha + \pi/2$ (thus $\vec{e}_{-\alpha} \cdot \vec{e}_{\tilde{\alpha}} = \vec{e}_{\alpha} \cdot \vec{e}_{-\tilde{\alpha}} = 0$) and $\vec{\eta} = (\eta, 0), \eta > 0$ is a fixed vector.

In the second lecture, we will describe in detail these directional wavelets and some of their properties. The following points will be developed:.

1. Generalities on directional wavelets: definition, examples.

2. Calibration of directional wavelets: scale and angular resolving power, benchmark tests.

3. Specific applications:

   - Directional filtering.
   - Fluid dynamics: visualization and measurement of a velocity field in a turbulent flow, disentangling of a wave train.

- Detection of symmetries: directional wavelets may be used for detecting (hidden) dilation-rotation symmetries in patterns, such as Penrose tilings, twisted fractals or the diffraction spectrum of a quasi-crystal. The tool here is the so-called *scale-angle measure* of the signal, namely the positive function

$$\mu_S(a,\theta) = \int |S(\vec{b},a,\theta)|^2 \, d^2\vec{b} \sim \int |\widehat{\psi}(ar_{-\theta}(\vec{k}))|^2 \, |\widehat{s}(\vec{k})|^2 \, d^2\vec{k}, \qquad (5)$$

  where $S(\vec{b},a,\theta)$ is the WT of the pattern s with a directional wavelet, usually a Cauchy wavelet. This includes an algorithm for testing whether all symmetries have been detected.

If time permits, a few indications will be given towards other generalizations of the CWT, such as 3-D wavelets, wavelets on a sphere, or time-dependent wavelets, as needed for analyzing motions, as in video sequences.

An introduction to the numerical solution of Stochastic Ordinary Differential Equations

by
K. Burrage

Department of Mathematics, University of Queensland, Brisbane, 4072, Australia
kb@maths.uq.edu.au

Abstract:

In recent years considerable attention has been paid to the numerical solution of stochastic ordinary differential equations (SODES), as SODEs are often more appropriate than their deterministic counterparts in many modelling situations. However, unlike the deterministic case numerical methods for SODEs are considerably less sophisticated due to the difficulty in representing the (possibly large number of) random variable approximations to the stochastic integrals.
Although Burrage and Burrage (1996) were able to construct strong order 1.5 stochastic Runge-Kutta methods for certain cases, in a more recent paper (Burrage and Burrage (1997)) it was shown that all known stochastic Runge-Kutta methods suffer an order reduction down to strong order 0.5 if there is non-commutativity between the functions associated with the multiple Wiener processes. This order reduction down to that of the Euler-Maruyama method imposes severe difficulties in obtaining meaningful solutions in a reasonable time frame, but these difficulties can be overcome by some new techniques involving Lie bracket evaluations.

An additional difficulty in solving SODEs arises even in the linear case, since it is not possible to write the solution analytically in terms of matrix exponentials unless there is again a commutativity property between the functions associated with the multiple Wiener processes. However, the work of Magnus (1954) (applied to deterministic non-commutative linear problems) can be applied to non-commutative linear SODEs and methods of strong order 1.5 for arbitrary, linear, non-commutative SODE systems can be constructed – hence giving an accurate approximation to the general linear problem.

Furthermore, for general nonlinear non-commutative systems with an arbitrary number of (d) Wiener processes it can shown that strong order 1 stochastic Runge-Kutta methods must evaluate a set of Lie brackets as well as the standard function evaluations and have at least d+1 stages. A method can then constructed which can be efficiently implemented in a parallel environment for this arbitrary number of Wiener processes.

This introductory talk will attempt to address these issues described above. In doing so, no prior knowledge of stochastic processes will be assumed.

Note: at

`http://www.cwi.nl/~jankok/kbpb4_ps.ps`

a background paper by K. Burrage and P.M. Burrage: *High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula* is available till one week after the conference.

Implementation issues in solving Stochastic Ordinary Differential Equations

by
K. Burrage

Department of Mathematics, University of Queensland, Brisbane, 4072, Australia
kb@maths.uq.edu.au

Abstract:

There are many issues that have to be addressed in developing an efficient implementation of a stochastic numerical method including the efficient and effective simulation of the stochastic integrals needed in the method formulation, the nature of the problem (non-commutative, stiff etc) and the selection of an appropriate method, as well as a means of providing effective variable stepsize strategies – which hitherto have been very poorly addressed. These issues will be discussed and some numerical results are presented which illustrate the efficacy of these new methods and techniques. At all times comparisons will be made with the deterministic situation in order to illustrate the relative paucity of suitable algorithms and codes in the stochastic case.

Wavelets and Adaptivity in Numerical Analysis – I, II

Wolfgang Dahmen, RWTH Aachen

The essence of multiscale techniques and wavelet concepts is the ability of separating effects associated with different scales of resolution. Moreover, significant coefficients in a wavelet expansion indicate the location and type of singularities. The mathematical foundation lies in cancellation properties of wavelet-type functions and the fact that in some range weighted sequence norms of expansion coefficients are equivalent to function norms of Sobolev or Besov type. This accounts for the potential of such concepts with regard to the analysis as well as to the efficient numerical treatment of problems involving the interaction of a wide range of scales.

Part I addresses the basic underlying concepts centering upon the above mentioned cancellation properties and norm equivalences, their background and main consequences in terms of preconditioning and matrix compression. These facts, which are exemplified in the context of some elliptic problems covering differential and singular integral operators, are also the main prerequisites for the design and analysis of adaptive techniques.

Part II is to focus on wavelet based adaptive techniques. First some applications of wavelet concepts within conventional discretizations in terms of finite element or finite volume schemes for transport dominated problems are outlined. This covers the design of problem adapted multigrid ingredients for convection diffusion equations with dominating convection or the acceleration of flux calculations for finite volume discretizations of conservation laws. In both cases adaptive refinements are based on monitoring quantities that can be viewed as wavelet coefficients. The performance of these schemes is illustrated by some numerical examples. The rest of the lecture is then concerned with a rigorous convergence analysis of such wavelet based adaptive schemes for a general class of elliptic problems again covering also operators of negative orders. In this setting it can be shown that a certain scheme provides *optimal* approximation rates. This means that whenever for a certain range of smoothness indices s the solution of the operator equation can be recovered with accuracy ϵ in the ideal case (i.e., with complete a-priori knowledge) by a linear combination of the order of $N(\epsilon) := \epsilon^{-1/s}$ wavelets then the adaptive algorithm produces an approximation using the same order of terms to achieve that accuracy. Moreover, the computational work can be

shown to remain proportional to order $N(\epsilon)$ as well. The approximability of the solution with the above order can be reinterpreted as *Besov regularity* of order ds (d the spatial dimension). The point here is that Besov regularity is a weaker scale than classical Sobolev regularity so that the approximation rate achieved by such an adaptive scheme is asymptotically better than that produced by ordinary uniform refinements precisely when the solution has deficient Sobolev regularity, e.g. as in the presence of reentrant corners for second order elliptic boundary value problems.

The underlying analysis brings up the following interesting points. In the course of the refinement process a well-quantified intermediate thresholding strategy (which actually results in intermediate coarsening) is essential for the claimed optimality. In this point the scheme differs from preceding ones. The numerical realization hinges on a new scheme for approximate matrix/vector multiplication suggested by the analysis that exploits not only the near sparseness of the stiffness matrices but also of the arrays of wavelet coefficients.

Inertial manifolds of parabolic pde's under time discretization

Jos van Dorsselaer
CWI
P.O. Box 94079
1090 GB Amsterdam,
The Netherlands
email: dorssela@cwi.nl

Abstract: Finite time error bounds may not lead to useful estimates when applied to time-stepping methods on long-time intervals. In order to analyse the qualitative behaviour of time-stepping methods in these cases, one has to proceed differently. For some parabolic equations the long-time behaviour is determined by invariant sets, such as periodic orbits, attractors and inertial manifolds (a finite dimensional set which attracts the solutions of a given partial differential equation exponentially). In this lecture we show that inertial manifolds can be approximated accurately for a large class of Runge-Kutta methods and BDF methods. As an application to the theory, the Ginzburg-Landau equation is considered.

References

F. Demengel & J.-M. Ghidaglia: Inertial manifolds for partial differential evolution equations under time-discretization: existence, convergence, and applications. *J. Math. Anal. Appl.* **155**, 177-225 (1991).

J.L.M. van Dorsselaer: Inertial manifolds under multistep discretization. Universität Tübingen, 1998.
Available from `http://na.uni-tuebingen.de/na/preprints.shtml`

J.L.M. van Dorsselaer & Ch. Lubich: Inertial manifolds of parabolic differential equations under higer-order discretizations. Universität Tübingen, 1998. To appear in *IMA J. Numer. Anal.*
Available from `http://na.uni-tuebingen.de/na/preprints.shtml`

C. Foias, G.R. Sell & R. Temam: Inertial manifolds for nonlinear evolutionary equations. *J. Diff. Eq.* **73**, 309-353 (1988).

D.A. Jones & A.M. Stuart: Attractive invariant manifolds under approximation. Inertial manifolds. *J. Diff. Eq.* **123**, 588-637 (1995).

T. Shardlow: Inertial manifolds and linear multi-step methods. *Numer. Algor.* **4**, 189-209 (1997).

A.M. Stuart: Perturbation theory for infinite dimensional dynamical systems, in *Theory and Numerics of Ordinary and Partial Differential Equations* (M. Ainsworth, J. Levesley, W.A. Light and M. Marletta, eds), 181-290. Oxford: Clarendon Press, 1995.

Deferred Correction with mono-implicit Runge-Kutta methods for first order IVPs

T. Van Hecke*

Vakgroep Toegepaste Wiskunde en Informatica, Universiteit Gent
Krijgslaan 281 - S9, B9000 Gent, Belgium

Keywords : Deferred Correction; Mono-implicit Runge-Kutta method; Stability
AMS classification : 65L05, 65L06, 65L20

Abstract

To reach a high order of accuracy for numerical solutions of IVPs with Mono-Implicit Runge-Kutta (MIRK) methods, the technique of deferred correction is used. Special attention is paid to the possible increase of the order and the stability of such schemes. Several schemes are given.

1 Introduction

For the numerical solution of first order IVPs

$$y' = f(x, y), \quad y(x_0) = y_0, \qquad y \in \mathbb{R}^n \text{ and } f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, \tag{1.1}$$

the following representation of s-stage implicit Runge-Kutta methods (IRK), known as *parameterized* IRK methods, was presented by Burrage *et al.* [1]:

$$\begin{aligned} y_{n+1} &= y_n + h \sum_{i=1}^{s} b_i f(x_n + c_i h, Y_i) \\ Y_i &= (1 - v_i) y_n + v_i y_{n+1} + h \sum_{j=1}^{s} x_{ij} f(x_n + c_j h, Y_j), \qquad i = 1, \ldots, s. \end{aligned}$$

Hence, a s-stage parameterized IRK method is completely determined by the tableau

$$\begin{array}{c|c|ccccc} c_1 & v_1 & x_{11} & x_{12} & \ldots & x_{1s} \\ c_2 & v_2 & x_{21} & x_{22} & \ldots & x_{2s} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_s & v_s & x_{s1} & x_{s2} & \ldots & x_{ss} \\ \hline & & b_1 & b_2 & \ldots & b_s \end{array} \tag{1.2}$$

Comparing this representation with the description of a general IRK method by means of its Butcher tableau (c, A, b) [2], it is easy to verify that the relationship $A = X + v.b^T$ holds. For all methods considered, we will assume that the row-sum condition holds, i.e. $A.e = c$ where e

*Research assistent of the University of Gent

is the s-vector with unit entries. By imposing that X (or X after a rearrangement of its rows and columns according to a same permutation) is a strictly lower triangular matrix one obtains mono-implicit Runge-Kutta (MIRK) methods [1].
Several results concerning MIRK methods have been established. Well-known are the following bounds : the order $p \leq s+1$ and the stage order is 3 at most. Also, in [1] a complete characterization is given of methods of order $p \leq 6$ with $s \leq p$ stages. Another family of MIRK methods is given in [7] : here $s = p$ and $c_i = 0, 1, \ldots, s-1$.
Also, there is no problem to find stable MIRK methods : when a MIRK method is applied to the test problem $y' = \lambda y$, $y(x_0) = y_0$ with fixed stepsize h, one obtains $y_n = R^n(\lambda h)\, y_0$ where $R(z) = P(e-v, z)/P(-v, z)$ with $P(w, z) = 1 + \sum_{i=1}^{s} z^i\, b^T.X^{i-1}.w$. This reveals one of the main problems one is confronted with when using MIRK methods : the Jacobian of the implicit system to be solved (which is of dimension n), is in practice approximated by the following non-linear expression in $J = \frac{\partial f}{\partial y}$:

$$I - \sum_{i=1}^{s} h^i\, J^i\, b^T.X^{i-1}.v\,.$$

This requires the computation of powers of J (an operation with complexity $\mathcal{O}(n^3)$). To avoid the computation of high powers of J, we propose to use the technique of *deferred correction* (DC). While Cash [3, 4] used this technique for BVPs, we will apply it for IVPs.

2 The DC algorithm

Suppose we want to approximate the solution of the IVP (1.1) on the mesh $x_0 < x_1 < x_2 < \ldots$ and let $h = \max_i h_i$ where $h_i := x_{i+1} - x_i$. Let Δy be the restriction of the continuous solution $y(x)$ to the grid and let η and η^* be an approximation to Δy.

We rely on a theorem proven by Skeel [6], which we reformulate in a slightly modified form.

Theorem 2.1 *Consider the DC scheme*

$$\begin{cases} \phi(\eta) = 0\,, \\ \phi(\eta^*) = \phi(\eta) - \phi^*(\eta)\,. \end{cases} \tag{2.3}$$

Suppose (i) $\eta = \Delta y + \mathcal{O}(h^p)$, *(ii)* $\psi(\Delta y) = \phi(\Delta y) + \mathcal{O}(h^{p^*})$, *and (iii)* $\psi(\Delta w) = \mathcal{O}(h^r)$ *for arbitrary functions* w *having at least* r *continuous derivatives, then*

$$\eta^* = \Delta y + \mathcal{O}(h^{\min(p^*, p+r)})\,. \tag{2.4}$$

As already mentioned, in our case ϕ will correspond to a MIRK method of order p and $\psi := \phi^* - \phi$ where ϕ^* corresponds to a MIRK method of order $p^* > p$ (we will systematically denote the quantities that relate to ϕ^* with a *-superscript : s^*, a^*_{ij}, b^*_i, c^*_i, $\ldots$). The interesting thing about ϕ^* being a MIRK method is that $\psi(\eta) = \phi^*(\eta)$ can be computed directly. Although one could argue that ϕ can be any RK method, we also choose it to be a MIRK method since in that case all systems to be solved have dimension n.

For ϕ we have $\phi(\Delta y)_n := \dfrac{1}{h_n}\,(y_{n+1} - y_n) - \sum_{i=1}^{s} b_i\, f(x_n + c_i\, h_n,\, Y_i)$, with $y_i := y(x_i)$ and

$$\begin{aligned} Y_i &= (1 - v_i)\, y_n + v_i\, y_{n+1} + h_n \sum_{j=1}^{i-1} x_{ij}\, f(x_n + c_j\, h_n,\, Y_j) \\ &= y_n + h_n \sum_{j=1}^{s} a_{ij}\, f(x_n + c_j\, h_n,\, Y_j) + \mathcal{O}(h_n^p)\,. \end{aligned}$$

The assumption (i) is a representation of the global error of the method ϕ with p the order of the method. If $y'(x) = f(x,\, y(x))$, then a Taylor series expansion gives

$$\phi(\Delta y)_n = (1 - b^T.e)\, f_n + \left(\frac{1}{2}(1 - 2\, b^T.A.e)\, f_n\, {}^y f_n + (1 - 2\, b^T.c)\, {}^x f_n\right) h_n + \mathcal{O}(h_n^2)$$

whereby the superscript denotes the derivation and the subscript n means that all evaluations are taken in $x = x_n$. One notices that, if the series expansion is carried out as far as $\mathcal{O}(h_n^p)$, in this way all the order conditions to achieve order p can be recognised. It thus becomes clear that the term in h_n^i, $0 \le i \le p - 1$ becomes zero when the method is of order p. We thus have $\phi(\Delta y) = \mathcal{O}(h_n^p)$. In the same way the condition (ii) of Theorem 2.1 expresses the order of the residual with the higher order method ϕ^*. Analogous to the previous derivation, $\phi^*(\Delta y)_n = \mathcal{O}(h_n^{p^*})$ can be deduced. The value r from assumption (iii) follows from the expansion of

$$\psi(\Delta w)_n = \sum_{i=1}^{s} b_i\, f(x_n + c_i\, h_n, Y_i) - \sum_{i=1}^{s^*} b_i^*\, f(x_n + c_i^*\, h_n, Y_i^*)\,.$$

One finds that $\psi(\Delta w) = \phi(\Delta w) - \phi^*(\Delta w)$ is $\mathcal{O}(h_n^r)$ where $r = \min(p,\, q)$ and

$$q = \begin{cases} 1 & \text{if } b^T.v \neq b^{*T}.v^* \\ 2 & \text{if } b^T.v = b^{*T}.v^* \\ & \text{but } |b^T.(c\,v) - b^{*T}.(c^*\,v^*)| + |b^T.(X\,v) - b^{*T}.(X^*.v^*)| + |b^T.v^2 - b^{*T}.v^{*2}| \neq 0 \\ 3 & \text{if } b^T.v = b^{*T}.v^* \\ & \text{and } |b^T.(c\,v) - b^{*T}.(c^*\,v^*)| + |b^T.(X\,v) - b^{*T}.(X^*.v^*)| + |b^T.v^2 - b^{*T}.v^{*2}| = 0 \\ & \text{but } \ldots \\ \ldots & \end{cases}$$

We thus find that, while the value r in condition (iii) is 1 in general, it can be raised to 2 or even higher. In [3, 4], where symmetric methods are used, the value $r = 2$ is obtained since for all symmetric methods $b^T.v = 1/2$. Combining the three conditions of Theorem 2.1, is it clear that there will be a gain $\mathcal{O}(h^g)$ with the DC technique based on ϕ and ϕ^*, where $g = \min(r,\, p^* - p) = \min(p,\, q, p^* - p)$. Since one may expect that, if $p = q = p^* - q$, the ratio *accuracy/computational cost* is optimal, we will call these schemes optimal.

The basis of coupling two methods by DC, can be enlarged to several methods. The general scheme of DC by coupling m methods is of the following form.

$$\begin{cases} \phi_1(\eta_1) &= 0, \\ \phi_1(\eta_i) &= \phi_1(\eta_{i-1}) - \phi_i(\eta_{i-1}), \qquad i = 2,\, \ldots,\, m \end{cases} \tag{2.5}$$

We will call ϕ_1 the basic method while ϕ_i, $i = 1,\, 2 \ldots,\, m$ are called the composing methods. Coupling several methods can be interesting for reasons of accuracy and/or stability. In this paper, we will restrict ourselves to schemes for which each method is used to raise the accuracy rather than to correct the stability properties.

3 Linear stability of DC-schemes

To analyze the linear stability properties of the method obtained, we introduce some new notations. Let $R_i(z) = N_i(z)/D_i(z)$ whereby $N_i(0) = D_i(0) = 1$ denote the linear stability function associated to method ϕ_i, then the linear stability function Z_m associated to the scheme (2.5) is recursively defined by

$$\begin{aligned} Z_1(z) &:= R_1(z), \\ Z_i(z) &:= \frac{(D_1(z) - D_i(z))\, Z_{i-1}(z) + N_i(z)}{D_1(z)}, \qquad i = 2, \ldots, m. \end{aligned}$$

In this way, it is clear that the denominator of $Z_m(z)$ is $D_1^m(z)$.

Several stability properties can be proven. A property which is useful in the construction of optimal DC schemes is given in the following theorem :

Theorem 3.1 *If $R_i(z) = \exp(z) + \mathcal{O}(z^{g\,i+1})$, $i = 1, \ldots, m$ then $Z_m(z) = \exp(z) + \mathcal{O}(z^{g\,m+1})$ if and only if $D_1(z) - D_i(z) = \mathcal{O}(z^g)$, $i = 1, \ldots, m$.*

If a DC scheme is set up consisting of m MIRK methods this condition means that, for $i = 0, 1 \ldots, g-1$, $b^T.X^i.v$ has the same value for all m methods.

We recall that our first aim is to reduce the computational work associated to the computation of high powers of J. Since the number of powers is determined by the degree of $D_1(z)$, we may want to choose a method ϕ_1 for which $D_1(z)$ is linear. In this respect, the trapezoidal rule looks very interesting since it is the only A-stable MIRK method for which $D_1(z)$ is linear which allows $g = 2$. Unfortunately, we have the following result :

Theorem 3.2 *The DC scheme (2.3) where ϕ is based on the trapezoidal rule and ϕ^* is a Runge-Kutta method M of order $p \geq 3$, cannot be A-stable.*

From the above result, it follows that if D_1 is linear, ϕ_1 can only be of first order if A-stability is required and thus only $g = 1$ is possible. If one looks for accurate A-stable schemes, it is thus necessary to consider schemes for which the denominator of the basic method is quadratic at least. In this case, it is still possible to avoid the computation of J^2 if D_1 is factorizable in linear terms. Then several systems (for which the iteration matrices are linear in J) have to be solved consecutively.

4 An Example

Case A : We select MIRK methods for which $c_i = i - 1$, $i = 1, 2, \ldots, s$. These methods, which still contain some parameters, are described in Section 3 of [7]. Since D_1 has to be quadratic at least, we look for a method ϕ_1 which is already of third order. It turns out that within the family considered it is possible to construct a L-stable fifth order method M_{345}, based on a 3 methods of orders 3, 4, and 5 respectively for which D_1 is factorizable and, if we call M_3 (resp. M_{34}) the method based on the third order (resp. third and fourth order) method alone, M_3 and M_{34} are A-stable. The values of the parameters to obtain this are $t = 2(\sqrt{3}+1)$ for $m = 3$, $t = 0$ and $s = 2(\sqrt{3}+1)$ for $m = 4$ and $s = -2 - 4\sqrt{3}/19$ and $t = 7/2 + 2\sqrt{3}$ for $m = 5$ (with stage order 3).

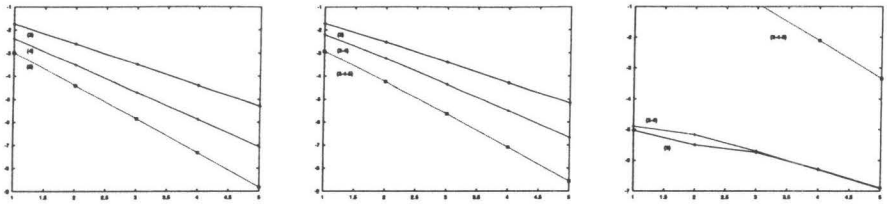


Figure 1: $-\log_2 h$ vs. $\log_{10}$ of the global error in $x = 1$ with (a) $\lambda = 0$ (left), (b) $\lambda = -1$ (middle) and (c) $\lambda = -1000$ (right) for the methods of case A.

As it is the case with RK methods in general, one can expect a possible order reduction when applying the method to stiff problems. Therefore, we apply the method to the Prothero-Robinson test problem [5]

$$y'(x) = \lambda\,(y(x) - g(x)) + g'(x), \qquad y(0) = g(0)\,, \tag{4.6}$$

with $g(x) = 10 - (10 + x)\,\exp(-x)$. We integrate this problem whose solution is $y(x) = g(x)$, up to $x = 1$ and we consider the global error for different values of the stiffness parameter λ and different values of the constant stepsize h. For $\lambda = 0$ the problem becomes explicit and the results obtained with deferred correction are those obtained with the last method used. The slopes of the lines in Figure 1 (a) confirm the theoretical order of the methods M_3, M_4, M_5.

For $\lambda \approx 0$, the problem is non-stiff and from Figure 1(b) one can easily deduce the expected order behaviour of the three methods M_3, M_{34} and M_{345}. However, as λ decreases, the behaviour changes. In Figure 1(c) we show the case where $\lambda = -1000$, in which case the problem is moderately stiff. One notices that M_{34} does not perform better than M_3, while M_{345} performs very badly. To understand the behaviour of the different schemes, we consider the LTEs and we look at the behaviour in the case $z = -\lambda h \to \infty$ and $h \to 0$ (this is what Prothero et al. call the stiff order).

5 The stiff order of DC schemes

When a parameterized RK method is applied to (4.6) with steplength h one obtains,

$$y_1 = \frac{(1 + \hat{h}\,B^T.(e - v))\,y + B^T.(h\,G'(0) - \hat{h}\,G(0))}{1 - \hat{h}\,B^T.v}\,. \tag{5.7}$$

where $\hat{h} := \lambda h$, $B^T := b^T.(I - \hat{h}\,X)^{-1}$ and $G(x)$ and $G'(x)$ are the s-vectors with entries $g(c_i\,x)$ and $g'(c_i\,x)$.

Theorem 5.1 *If a parameterized RK method of order p with stage order $q \le p$ is applied to (4.6), then*

$$y(h) - y_1 = \frac{h^{q+1}}{(q+1)!}\,C_{q+1}(\hat{h})\,y^{(q+1)}(0) + \mathcal{O}(h^{q+2})\,, \tag{5.8}$$

where $C_{q+1}(\hat{h}) = 1 - (q+1)\,b^T.c^q + \dfrac{\hat{h}B^T.(c^{q+1} - (q+1)\,A.c^q)}{1 - \hat{h}\,B^T.v}$.

If a method is fitted to solve stiff problems, the rational function $C(z) \sim z^{-p_z}$ with $p_z \geq 0$ as $z \to \infty$. For DC-schemes we need to know how the corresponding expression grows out of the expressions for the composing methods. Therefore, we define $S(h,\, \hat{h}) := B^T.(h\, G'(0) - \hat{h}\, G(0))$, such that we obtain from (5.7) that $y_1 = [N(\hat{h})\, y + S(h,\, \hat{h})]/D(\hat{h})$. When the scheme (2.5) is applied to problem (4.6), one obtains the approximations $\tilde{y}_{1,i} = Z_i(\hat{h})\, y + W_i(h, \hat{h})$, $i = 1,\, 2 \ldots,\, m$, where

$$Z_i(z) := \frac{(D_1(z) - D_i(z))\, Z_{i-1}(z) + N_i(z)}{D_1(z)}, \qquad Z_1(z) := \frac{N_1(z)}{D_1(z)},$$
$$W_i(z,\, \hat{z}) := \frac{(D_1(z) - D_i(z))\, W_{i-1}(z,\, \hat{z}) + S_i(z,\, \hat{z})}{D_1(z)}, \quad W_1(z,\, \hat{z}) := \frac{S_1(z,\, \hat{z})}{D_1(z)}.$$

If we now consider the case where $h \to 0$ and $\hat{h} \to \infty$ and we define $\tilde{q}_m := \min\limits_{1 \leq i \leq m} \{q_i | C_{q_i+1}(\hat{h}) \neq 0\}$ where q_i and $C_{i,q_i+1}(\hat{h})$ follow from (5.8) for method ϕ_i, then

$$y(h) - \tilde{y}_{1,m} = \frac{h^{\tilde{q}_m+1}}{(\tilde{q}_m+1)!} y^{(\tilde{q}_m+1)}(0)\, \tilde{C}_{m,\tilde{q}_m+1}(\hat{h}) + \mathcal{O}(h^{\tilde{q}_m+2}),$$

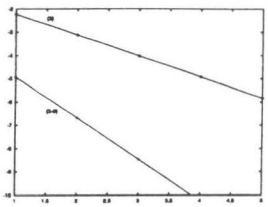
where $\tilde{C}_{1,\tilde{q}_m+1}(z) := C_{1,\tilde{q}_m+1}(z)$ and

$$\tilde{C}_{i,\tilde{q}_m+1}(z) := \frac{(D_1(z) - D_i(z))\, \tilde{C}_{i-1,\tilde{q}_m+1}(z) + D_i(z)\, C_{i,\tilde{q}_m+1}(z)}{D_1(z)}.$$

If we now return to Case A, we find that $\tilde{q}_3 = 2$ since $q_1 = 2$ and $q_2 = q_3 = 3$ and from which one easily finds that that for $z \to \infty$ $\tilde{C}_{1,3} \sim z^{-1}$ and $\tilde{C}_{2,3} \sim z^{-1}$ but $\tilde{C}_{3,3} \sim z^1$.

Case B : A third and last example illustrates the possibility to have a stable DC-scheme with gain $g = 3$ with a stable s_1-stage method of order 3 and a s_2-stage method of order 6 who both have the maximum stage-order 3. To construct this scheme, we first examined the cases where the total number of stages $s_1 + s_2$ is minimal, taking into account that $s_2 \geq 5$ to obtain order 6 and and $s_1 \geq 3$ to obtain order 3 and stage order 3 and we made use of the fact that expressions of the form $b^T.X^i.v$ and $b^T.X^i.e$ are connected to each other by the order equations. This technique showed that it was impossible to have A-stability for $s_1 = 3$ and $s_2 = 5$ or $s_2 = 6$. We thus chose $s_1 = 4$ and $s_2 = 5$. For the sixth order method, we used the family in [1]. This family contains 2 parameters $c_3^{(6)}$ and $c_4^{(6)}$. A family of third order methods with 4 stages which has stage order 3 and for which the denominator of the stability function has fixed linear and quadratic coefficients also contains 2 parameters $c_3^{(3)}$ and $c_4^{(3)}$, whereby stability requires that $|(c_3^{(3)} - 1)/c_3^{(3)}| < 1$. There is one possibility, $c_4^{(6)} = 1 - c_3^{(6)}$, to make the DC scheme A-stable and L-stability can be obtained for $c_3^{(3)} = \sqrt{2}$. Considering the $\tilde{C}$-expressions reveals that both the basic method and the DC-scheme are $\sim z^{-1}$ irrespective of the choice made for $c_3^{(3)}$. The two remaining conditions, which express that $b^T.(cv)$ and $b^T.(v^2)$ should have a fixed value for both methods, are then used to determine $c_3^{(6)}$ and $c_4^{(3)}$. We mention the following solution :

0	0				
1	1	0			
.7071067812	.7928932188	.0606601718	−.1464466094		
−.2670411948	−.2606042131	−.2932210260	−.1257432186	.4125272629	
		1.0863664648	.3492484895	.0353379297	−.4709528840



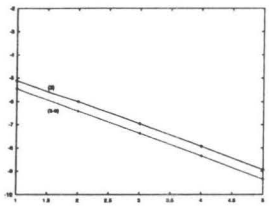


Figure 2: $-\log_2 h$ vs. $\log_{10}$ of the global error in $x = 1$ with (a) $\lambda = -1$ (left) and (b) $\lambda = -1000$ (right) for the methods of case B.

0	0					
1	1	0				
$c_3^{(6)}$	$v_3^{(6)}$	−1.249716874	−.4341220130			
$1 - c_3^{(6)}$	$1 - v_3^{(6)}$	.4341220130	1.249716874	0		
.5	.5	.1410807817	−.1410807817	−.0077889892	.0077889892	
		.1871016491	.1871016491	−.0047942664	−.0047942664	.6353852345

where $c_3^{(6)} = -.5322765429$ and $v_3^{(6)} = 1.151562344$.

A final analysis shows that, apart from small regions of instability along the imaginary axis, the basic method of order three is A-stable and the DC-scheme itself is L-stable.

6 Conclusion

High order DC schemes can be constructed, but that it is insufficient to consider only linear stability. One can make sure that the stability of the DC scheme is ensured also for non-linear systems of equations. For the non-stiff case, there is a natural mechanism present in the DC scheme to perform error control and stepsize selection. But for the stiff case, this mechanism is no longer present due to order reduction.

References

[1] K. Burrage, F.H. Chipman and P.H. Muir, Order results for mono-implicit Runge-Kutta methods, *SIAM J. Numer. Anal.* **31** (1994) 876–891.

[2] J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations : Runge-Kutta and General Linear Methods*, J. Wiley, Chichester, (1987).

[3] J.R. Cash, A variable order deferred correction algorithm for the numerical solution of nonlinear two point boundary value problems, *Comp. & Maths. with Appls.* **9** (1983) 257–265.

[4] J.R. Cash and H.H.M. Silva, Iterated deferred correction for linear two-point boundary value problems, *Comp. Appl. Math.* **15** (1996) 55–75.

[5] A. Prothero and A. Robinson, On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations, *Mathematics of Computations* **28** (1974) 145–162.

[6] R.D. Skeel, A theoretical framework for proving accuracy results for deferred correction, *SIAM J. Numer. Anal.* **19** (1981) 171–196.

[7] M. Van Daele, G. Vanden Berghe and H. De Meyer, A general theory of stabilized extended one-step methods for ODEs, *Intern. J. Computer Math.* **60** (1996) 253-263.

The convergence of Runge–Kutta methods for delay differential equations

Karel in 't Hout
Mathematical Institute, Leiden University

In this talk we consider the numerical solution of initial value problems for delay differential equations,

$$(1) \quad U'(t) = f(t, U(t), U(t-\tau)) \quad (t > 0), \quad U(t) = g(t) \quad (-\tau \leq t \leq 0),$$

where f, g denote given (vector-valued) functions, τ denotes a given, fixed, positive real number, and $U(t)$ (for $t > 0$) is unknown. Initial value problems of the type (1) arise in many branches of science and engineering, such as physiology, epidemiology, and electrical circuit simulation. A popular approach to obtain numerical (step-by-step) methods for (1) consists of the adaptation of known step-by-step methods for the numerical solution of initial value problems (1) without a delay argument $U(t-\tau)$. The adaptation to general problems (1) is done by means of an interpolation procedure, which computes, in each (time-)step of the numerical process, approximations to the exact solution of (1) at a certain number of previous time-points t.

In this talk we shall consider the class of numerical step-by-step methods for (1) that is obtained by adaptation of the well-known class of Runge–Kutta methods (see e.g. [1]) using the interpolation procedure that has been introduced in [2]. We are interested in the convergence behaviour of this class of methods in the numerical solution of general, *non-stiff* initial value problems (1). In [1] the result was obtained that if the stepsizes are *constant* and equal to an integer fraction of the delay τ, then any given method from the class under consideration has order of convergence p, where p is the order (of consistency) of the underlying Runge–Kutta method. Up to now the important problem has been completely open, however, whether for any given method under consideration the same (high) order of convergence holds for general cases of *variable* stepsize sequences, whenever the number of support points for the interpolation procedure is sufficiently large. (We note that in the case of [1] there is in fact no interpolation error.) In this talk we will present a main result on this problem, which substantially extends the result obtained in [1]. We will illustrate our (convergence) result by various numerical experiments.

References

[1] E. Hairer, S.P. Nørsett & G. Wanner: *Solving Ordinary Differential Equations I. Nonstiff Problems.* Springer-Verlag, Berlin, 2nd ed. (1993).

[2] K.J. in 't Hout: *A new interpolation procedure for adapting Runge–Kutta methods to delay differential equations.* BIT **32**, 634–649 (1992).

Talk I: Multilevel Frames and Space Splittings with Applications to Iterative Methods

Talk II: Multilevel Discretization Schemes for the Single Layer Potential Equation

Peter Oswald

Bell Labs, Lucent Technologies
600 Mountain Av., Rm. 2C-403
Murray Hill, NJ 07974-0636

e-mail: poswald@research.bell-labs.com
www: http://cm.bell-labs.com/who/poswald

Abstract

This is the support material for two talks given at the 1998 Dutch Conference of Numerical Mathematicians. In Talk I (sections 1-3 below), we review the notions of frames and stable space splittings in a Hilbert space setting. While the frame concept was developed as part of (non-harmonic) Fourier analysis, mainly in connection with signal processing applications, the latter theory of stable subspace splittings has led to a better understanding of iterative solvers (multigrid/multilevel resp. domain decomposition methods) for large-scale discretizations of symmetric elliptic variational problems in Sobolev spaces. In Talk II (see section 4), several aspects (apriori and aposteriori compression, preconditioning, resolution of singularities) of solving the single layer potential equation by multiscale methods are discussed. Although our analysis is restricted to the unit square $[0,1]^2$, some observations generalize, and are worth further investigation.

1 Frames

The notion of a *frame* in a Hilbert space V was introduced in [12]. A first survey with emphasis on frames was [16], see also [5, Chapter 3], [9, Chapter 3]. A more recent and comprehensive source is the collection [29] which we recommend for further reading.

Definition 1 *Let $F \equiv \{f_k\}$ be an at most countable system of elements in V.*
a) *F is a* **frame** *in V if there are two constants $0 < A \le B < \infty$ such that*

$$A\|f\|^2 \le \sum_k |(f, f_k)|^2 \le B\|f\|^2 \qquad \forall f \in V\,. \tag{1}$$

The optimal constants A, B in (1) are the lower and upper frame bounds, respectively, their ratio B/A defines the condition of F and will be denoted by $\kappa(F)$.
b) *F is a* **Riesz basis** *in V if F is dense in V and there are constants $0 < \tilde{A} \le \tilde{B} < \infty$ such that*

$$\tilde{A}\|f\|^2 \le \sum_k c_k^2 \le \tilde{B}\|f\|^2 \qquad \forall f = \sum_k c_k f_k\,. \tag{2}$$

Assuming **b)**, it can indeed be proved that any $f \in V$ possesses a ***unique*** V–converging series representation

$$f = \sum_k c_k f_k \ , \quad c = (c_k) \in \ell^2 \ . \tag{3}$$

Any complete orthonormal systems $\{e_j\} \subset V$ (CONS) is obviously both a frame and a Riesz basis. Any Riesz basis is a frame (with $A = 1/\tilde{B}$, $B = 1/\tilde{A}$). This follows from comparing (2) with the following result:

Theorem 2 [12] *A system F satifies (1) (i.e., is a frame) if and only if*

$$B^{-1}\|f\|^2 \leq |||f|||^2 \equiv \inf_{c:\, f=\sum_k c_k f_k} \|c\|_{\ell^2}^2 \leq A^{-1}\|f\|^2 \quad \forall f \in V \ . \tag{4}$$

For (4) to hold, it is implicitly required that any $f \in V$ possesses at least one V–converging series representation (3). Unless a frame F is a Riesz basis, such a series is ***nonunique***. Nevertheless, frames can be used to represent elements from V in essentially the same way as CONS or Riesz bases. To produce a representation formula, we need some standard definitions. Let F be a frame in V. Then the ***synthesis operator*** R given by

$$R \ : \ c = (c_k) \in \ell^2 \quad \longmapsto \quad Rc = \sum_k c_k f_k \in V$$

is well-defined and bounded. Its adjoint $R^* \ : \ V \to \ell^2$ takes the form

$$f \in V \quad \longmapsto \quad R^* f = ((f, f_k)) \in \ell^2$$

and is called ***analysis operator***. The boundedness of R and R^* follows exclusively from the upper estimate in the definition (1). The two-sided inequality (1) can be rephrased as

$$A(f,f) \leq \|R^* f\|_{\ell^2}^2 = (RR^* f, f) \leq B(f,f) \qquad \forall f \in V \ ,$$

which shows that the *frame operator* $\mathcal{P} = RR^* \ : \ V \to V$ is symmetric and has a bounded inverse:

$$\mathcal{P} = \mathcal{P}^* \ , \quad A\,\mathrm{Id} \leq \mathcal{P} \leq B\,\mathrm{Id} \ , \ \|\mathcal{P}\|_{V\to V} = B \ , \quad \frac{1}{B}\mathrm{Id} \leq \mathcal{P}^{-1} \leq \frac{1}{A}\mathrm{Id} \ , \ \|\mathcal{P}^{-1}\|_{V\to V} = \frac{1}{A} \ .$$

Here, A, B are the frame bounds of F. As a consequence, the spectral condition number of $\mathcal{P}$ coincides with the frame condition:

$$\kappa(\mathcal{P}) = \|\mathcal{P}\|_{V\to V}\|\mathcal{P}^{-1}\|_{V\to V} = \kappa(F) \ .$$

Obviously,

$$f = \sum_k (\mathcal{P}^{-1} f, f_k) f_k = \sum_k (f, \mathcal{P}^{-1} f_k) f_k \qquad \forall f \in V \ . \tag{5}$$

The system $\tilde{F} = \{\tilde{f}_k = \mathcal{P}^{-1} f_k\}$ is called ***dual frame***. It is easy to see that $\tilde{F}$ is indeed a frame, with frame operator $\mathcal{P}^{-1}$. Finally, note that there is another interesting operator $\tilde{\mathcal{P}} = R^* R \ : \ \ell^2 \to \ell^2$ which is also symmetric (in ℓ^2) but not necessarily invertible. Its matrix representation $\tilde{\mathcal{P}} = (\tilde{\mathcal{P}}_{k,j} = (f_j, f_k))$ suggests the name *Gramian of* F for $\tilde{\mathcal{P}}$. $\tilde{\mathcal{P}}$ can also be used for characterizing properties of a frame.

(5) is the desired canonical decomposition-reconstruction formula. It even gives the ***best*** representation (3) with respect to F such that the infimum in (4) is achieved. Its practical use requires, in one way or the other, to compute $\mathcal{P}^{-1}$ on certain elements of V, or equivalently, to solve the operator equation

$$\mathcal{P} g = h$$

for given $h \in V$. It was already proposed in [12] (see also [30, Section 8.2]) that ***Richardson iteration***

$$g^{(n+1)} = g^{(n)} + \omega(h - \mathcal{P} g^{(n)}) \ , \quad n \geq 0 \ , g^{(0)} \in V \ , \quad \omega = 2/(A+B) \ , \tag{6}$$

could be used. The convergence rate of the iteration (6) is given by

$$\rho_R = \rho(\mathrm{Id} - \omega\mathcal{P}) = 1 - \frac{2}{1+\kappa(F)} ,$$

it exclusively depends on the frame condition. Since $\mathcal{P}$ is symmetric, Richardson iteration can be replaced by the *conjugate gradient method* which would result in an even better convergence rate and avoid knowledge of good bounds for A, B. Other iterative methods might be tried as well.

There is another tricky point. In many applications, the theoretical investigations are for infinite frames (in infinite-dimensional V) while the algorithms work with sections $F_n = \{f_1, \ldots, f_n\}$ of the frame. An example from [4] shows that one should be cautious. Let

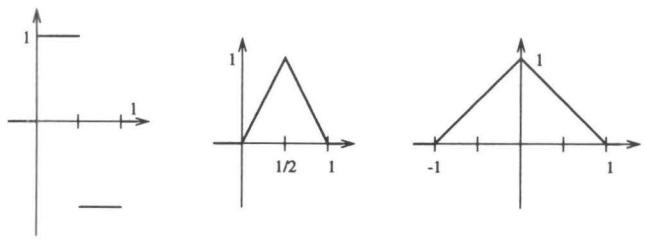
$$F = \{f_1 = e_1, f_2 = e_1 + e_2/2, \ldots, f_k = e_{k-1} + e_k/k, \ldots\} ,$$

where $\{e_k\}$ is a CONS in V. F is a frame. However, if one considers its sections F_n as frames in the subspaces $\operatorname{span} F_n \subset V$, then the corresponding lower frame bounds A_n deteriorate as $n \to \infty$. It can be shown that $\kappa(F_n) \geq (n!)^2$. Thus, working with the sections F_n of a frame F needs special care. In contrast, if F is a Riesz basis then the inequalities in (2) are automatically preserved for any subsystem, with the same (or better) constants, which yields $\kappa(F_n) \leq \kappa(F)$.

Frames have been considered mainly in connection with image and signal processing applications. The most prominent investigations are connected with **irregular sampling** ([12],[30, Chapter 8]), **Gabor frames** ([30, Chapter 3 and 7],[14],[13]), and **multilevel systems** originating from some kind of multiresolution analysis of subspaces $\{V_j\}$. E.g., given a nonzero function $\psi \in L_2(\mathbb{R})$, we can define *wavelet-like systems* by using integer shifts and dyadic dilation:

$$F_\psi = \{\psi_{j,i}(t) = 2^{j/2}\psi(2^j t - i) , \quad j, i \in \mathbb{Z} \}.$$

More information can be found in [9, 5, 29, 30]. The classical counterparts of this construction are the Haar and the Faber-Schauder system. These are obtained if the functions ψ depicted in Figure 1 a) and b) are used, respectively. Both choices lead to linearly independent systems. In the Haar case, the resulting wavelet system F_ψ is even a CONS in $L_2(\mathbb{R})$. Moreover, after suitable scaling it is a Riesz basis for the Sobolev spaces $H^s(\mathbb{R})$ with $-1/2 < s < 1/2$. The Faber-Schauder system associated with the ψ in Figure 1 b) is not a Riesz basis in $L_2(\mathbb{R})$ but in $H^s(\mathbb{R})$, $1/2 < s < 3/2$. The system F_ψ resulting from the hat function in Figure 1 c) leads to a system which contains redundancy, and yields frames (not Riesz bases) in $H^s(\mathbb{R})$ if $0 < s < 3/2$. All these systems have generalizations to higher dimensions and to the finite element setting on bounded domains. See [22, 6, 21, 8] for further results.



a) Haar wavelet

b) Faber-Schauder 'wavelet' (hierarchical basis)

c) BPX 'wavelet' (multilevel nodal basis frame)

Figure 1: Three basic 'wavelets' ψ

2 Stable space splittings

We come to the concept of stable space splittings which originated from work on the theoretical foundation of fictitious domain, domain decomposition methods, and multigrid methods [11, 32] for variational problems. It generalizes the frame concept in two directions: The individual f_k are replaced by Hilbert spaces V_j, and the assumption $f_k \in V$ is relaxed (instead of requiring $V_j \subset V$ we only assume the existence of suitable mappings $R_j : V_j \to V$, the scalar product on V_j need not be inherited from V). This allows for a broader range of applications to be covered (outer approximation schemes, block-iterative schemes, etc.). However, many basic ideas remain the same (this will be expressed by the notation used below).

Again, let V be the basic Hilbert space, with $(\cdot,\cdot)$ resp. $\langle\cdot,\cdot\rangle \equiv \langle\cdot,\cdot\rangle_{V'\times V}$ as basic scalar product resp. duality pairing. Consider a *symmetric V-elliptic variational problem*

$$u \in V : \quad a(u,v) = \langle f, v\rangle \qquad \forall v \in V \tag{7}$$

to be solved. (7) is equivalent to the operator equation $Au = f$, where $f \in V'$ and $A : V \to V'$ is defined by $\langle Au, v\rangle = a(u,v)$. Since symmetry and V-ellipticity require *symmetry, continuity*, and *coercivity* of the bilinear form $a(\cdot,\cdot)$, $\{V; a(\cdot,\cdot)\}$ (i.e., the space V equipped with the scalar product $a(\cdot,\cdot)$) is an isomorphic copy of $\{V;(\cdot,\cdot)\}$.

Let V_j, $j = 1, 2, \ldots$, be an at most countable family of Hilbert spaces, with $(\cdot,\cdot)_j, \langle\cdot,\cdot\rangle_j$ introduced similarly. To each V_j we assign its own symmetric V_j-elliptic bilinear form $b_j(\cdot,\cdot) : V_j \times V_j \to \mathbf{R}$ which in particular means that $\{V_j; b_j(\cdot,\cdot)\}$ are Hilbert spaces. The V_j and $b_j(\cdot,\cdot)$ will be used to create auxiliary variational problems, and to compose from their solution operators an approximate inverse to A. The latter is then used as a preconditioner in an iterative method for solving (7), see section 3 for the details. It is not assumed that the V_j are subspaces of V (but it is implicit that they correspond to certain portions of V, see below).

Denote the Hilbert sum of this family by $\tilde{V}$, i.e., for

$$\tilde{u} = (u_j),\ \tilde{v} = (v_j), \quad u_j, v_j \in V_j \ \ \forall j$$

set

$$\tilde{a}(\tilde{u},\tilde{v}) = \sum_j b_j(u_j, v_j)$$

which makes sense as a scalar product on

$$\tilde{V} = \{\tilde{u} : \tilde{a}(\tilde{u},\tilde{u}) < \infty\}\,.$$

Finally, consider bounded linear mappings $R_j : V_j \to V$. Formally, they can be considered as the components of an operator $R : \tilde{V} \to V$ given by $R\tilde{u} = \sum_j R_j u_j$.

Definition 3 *([22]) The system $\{\{V_j; b_j\}, R_j\}$ gives rise to a* **stable splitting** *of $\{V; a\}$ which will be expressed by the short-hand notation*

$$\{V; a\} \cong \sum_j R_j\{V_j; b_j\}\,, \tag{8}$$

if there are two constants $0 < \tilde{A} \le \tilde{B} < \infty$ such that

$$\tilde{A}a(u,u) \le |||u|||^2 \equiv \inf_{\tilde{u}\in\tilde{V}:\, u=R\tilde{u}} \tilde{a}(\tilde{u},\tilde{u}) \le \tilde{B}a(u,u) \quad \forall\, u \in V\,. \tag{9}$$

The optimal constants $\tilde{A}, \tilde{B}$ in (9) will be called lower and upper stability constants, and their ratio $\kappa = \tilde{B}/\tilde{A}$ condition of the splitting (8).

It should be noted that (9) implicitly requires that R makes sense (convergence of the sum if infinitely many V_j are involved) and yields a bounded operator from $\tilde{V}$ *onto* V, i.e., $\mathrm{ran}(R) = V$. The similarity of this definition with (4) in Theorem 2 is obvious. The adjoint $R^* : V \to \tilde{V}$ is defined as

$$R^* : u \in V \quad \longmapsto R^*u = (R_1^*u, R_2^*u, \ldots) \in \tilde{V} ,$$

where the components $R_j^* : V \to V_j$ are determined by solving the auxiliary variational problems:

$$b_j(R_j^*u, v_j) = a(u, R_j v_j) \qquad \forall\, v_j \in V_j . \tag{10}$$

Introduce the bounded linear operators

$$\mathcal{P} = RR^* : u \in V \longmapsto \mathcal{P}u = \sum_j T_j u \in V \quad (T_j = R_j R_j^* : V \to V) \tag{11}$$

and

$$\tilde{\mathcal{P}} = R^*R : \tilde{u} \to \tilde{\mathcal{P}}\tilde{u} \in \tilde{V} \tag{12}$$

where $\tilde{\mathcal{P}}$ can be considered as operator matrix with entries $\tilde{P}_{jk} = R_j^* R_k$. Following some tradition [11, 32], $\mathcal{P}$ is called *Schwarz operator* associated with the stable splitting (8) while the operator matrix associated with $\tilde{\mathcal{P}}$ will be called *extended Schwarz operator* (it is nothing but the generalization of the Gramian for frames discussed in Section 1, and the abstract analog of the matrix of the semi-definite system [17]).

Theorem 4 *The Schwarz operator (11) associated with a stable splitting (8) is symmetric positive definite and has a bounded inverse. Moreover,*

$$|||u|||^2 = a(\mathcal{P}^{-1}u, u) \qquad \forall\, u \in V ,$$

and

$$\frac{1}{\tilde{B}}\mathrm{Id} \le \mathcal{P} \le \frac{1}{\tilde{A}}\mathrm{Id} , \quad \tilde{A}\mathrm{Id} \le \mathcal{P}^{-1} \le \tilde{B}\mathrm{Id} , \qquad \kappa(\mathcal{P}) = \kappa .$$

With $\phi = R\tilde{\phi} \in V$, $\tilde{\phi} \in \tilde{V}$ defined from f in an appropriate way, $u \in V$ solves the variational problem (7) if it solves the operator equation

$$\mathcal{P}u = \phi , \tag{13}$$

or, equivalently, $u = R\tilde{u}$ for any solution $\tilde{u} \in \tilde{V}$ of the operator equation

$$\tilde{\mathcal{P}}\tilde{u} = \tilde{\phi} . \tag{14}$$

The computational aspect of these reformulations of (7) will be discussed in section 3. Standard examples of stable space splittings are discussed in [22, 26, 27]. A particularly important example, with deep connections to approximation, function space and interpolation theory, are **multilevel splittings** associated with a hierarchy of spaces

$$V_0 \to V_1 \to \ldots \to V_{j-1} \to V_j \to \ldots \to V , \tag{15}$$

where the relation $V_{j-1} \to V_j$ is described by an *embedding operator* $I_j : V_{j-1} \to V_j$ (these I_j are also called *prolongations* or *intergrid transfer operators*). If the spaces are *nested*, i.e., if $V_{j-1} \subset V_j$, then the natural embeddings are often the prefered choice. Define $R_j^J = I_J \ldots I_{j-1} : V_j \to V_J$, $0 \le j \le J$, and $R_j = \lim_{J\to\infty} R_j^J : V \to V$ (the existence of the latter operators needs verification). The multilevel splittings of interest are

$$\{V_J; a_J\} \cong \sum_{j=0}^{J} R_j^J\{V_j; b_j\} , \qquad \{V; a\} \cong \sum_{j=0}^{\infty} R_j\{V_j; b_j\} . \tag{16}$$

In applications to differential and integral equations on a domain Ω, where (7) is related to energy minimization in Sobolev norms, the b_j are given by scaled L_2–scalar products. Then the verification of the stability of the splittings in (16) can be reduced to the study of *Jackson-Bernstein inequalities* and *approximation spaces* associated with (15) [2, 22, 6]. Other techniques (e.g., using information on strengthened Cauchy-Schwarz inequalities

$$a(R_j u_j, R_l u_l) \leq \gamma_{jl} b_j(u_j, u_j)^{1/2} b_l(u_l, u_l)^{1/2} \qquad \forall\, u_j \in V_j\,,\ u_l \in V_l\ ) \tag{17}$$

can be found in [32, 34, 1].

We can introduce some general operations on stable space splittings which allow us to modify a given one, in order to adapt it to a specific application or to optimize the implementation with respect to a given hardware platform. In [18, 22], we described *refinement* (replace some of the components $\{V_j; b_j\}$ of a splitting by stable splittings of their own), *clustering* (the inverse operation), and *selection* (replace some V_j by subspaces $\hat{V}_j \subset V_j$ or drop some components; this operation corresponds to selecting subsystems of a frame, and may lead to a detoriation of the condition number of the splitting). Furthermore, *tensor-product techniques* can be explored [19] to obtain splittings for higher-dimensional applications. Another variation is to consider *mappings of stable splittings* to produce stable splittings for the range of a certain operator $T : V \to \hat{V} = \mathrm{ran}(T)$. This applies, e.g., to problems associated with trace spaces [25].

3 Iterative solvers

We come to some consequences of the notion of stable space splittings for the construction of iterative solution methods for solving the variational problem (7). Theorem 4 provides the tools. Assume that the splitting (8) is stable, and recall that

$$a(u, v) = \langle Au, v\rangle\ , \quad b_j(u_j, v_j) = \langle B_j u_j, v_j\rangle_j\ ,$$

defines invertible operators $A : V \to V'$, $B_j : V_j \to V_j'$. Introduce the dual operators $R_j' : V' \to V_j'$ by

$$\langle R_j' f, v_j\rangle_j = \langle f, R_j v_j\rangle \qquad \forall\, f \in V',\ v_j \in V_j\ .$$

It follows that $R_j^* = B_j^{-1} R_j' A$. Note that $\tilde{\phi}_j = B_j^{-1} R_j' f$ is the right choice in (13-14). Thus, the additive Schwarz operator has the representation $\mathcal{P} = CA$, where

$$C = \sum_j R_j B_j^{-1} R_j' \equiv \sum_j \hat{T}_j\ :\ V' \to V \tag{18}$$

satisfies $C' = C$. Obviously, C can be considered as a (symmetric) *preconditioner* or *approximate inverse* for A, and the switch from (7) to the equivalent formulation (13) as a *preconditioning method*, the quality of which critically depends on the condition κ of the splitting.

Following this setup, several iterative methods for solving (7) based on auxiliary subproblems associated with the given stable splitting can be introduced and analyzed (see [32, 34, 1, 22, 18]). To be practical, consider a *finite stable splitting*

$$\{V; a\} \cong \sum_{j=0}^{J} R_j\{V_j; b_j\} \qquad (\dim V_j = N_j < \infty\,,\ \dim V = N < \infty\,)\,. \tag{19}$$

The following basic algorithms associated with (19) have been formulated in [32]:

(AS) Additive Schwarz method. Starting with an initial guess $u^{(0)} \in V$, repeat

$$u^{(n+1)} = u^{(n)} + \omega \sum_{j=1}^{J} \hat{T}_j(f - Au^{(n)})\ ,$$

until a stopping criteria is satisfied.

(MS) Multiplicative Schwarz method. Starting with an initial guess $u^{(0)} \in V$, repeat

$$\begin{aligned} v^{(J+1)} &= u^{(n)} , \\ v^{(j)} &= v^{(j+1)} + \omega \hat{T}_j (f - A v^{(j+1)}) , \quad j = J, \ldots, 0 , \\ u^{(n+1)} &= v^{(0)} , \end{aligned}$$

until a stopping criteria is satisfied.

Variations (pcg-iterations, symmetric multiplicative methods) are possible. Note that the ordering of the subproblems has impact only on the multiplicative method **(MS)**. The *relaxation parameter* $\omega > 0$ can be used to properly *scale* the subproblems, and to enhance the convergence behavior. A special case of **(AS)** is the iteration (6) mentioned in connection with frame decompositions.

An elegant way to analyze the above iterations is to rewrite them in terms of classical iterative methods applied to the operator matrix $\tilde{\mathcal{P}}$ (which is now of size $J+1$) as proposed in [17, 18]. Richardson iteration and the SOR-method applied to the "matrix" problem (14) in $\tilde{V}$ transform into the iterations **(AS)** and **(MS)** in V, respectively, if the mapping R is applied. This leads to the following convergence result:

Theorem 5 *Let (19) be a finite stable space splitting, with stability constants $\tilde{A}, \tilde{B}$, and condition κ.*
a) *The additive method* **(AS)** *converges for $0 < \omega < 2\tilde{A}$. The optimal convergence rate is achieved for $\omega^* = 2\tilde{A}\tilde{B}/(\tilde{A} + \tilde{B})$:*

$$\rho_{AS}^* = \inf_{0<\omega<2/\tilde{A}} \rho_{AS,\omega} = 1 - \frac{2}{1+\kappa} . \tag{20}$$

b) *For the multiplicative algorithm* **(MS)**, *convergence is guaranteed if $0 < \omega < 2/\gamma$, where $\gamma = \max_j \gamma_{jj} \leq 1/\tilde{A}$ (the γ_{jj} are defined in (17)). The optimal convergence rate can be estimated by*

$$(\rho_{MS}^*)^2 = \inf_{0<\omega<2/\gamma} (\rho_{MS,\omega}^*)^2 \leq 1 - \frac{1}{\log_2(4(J+1)) \cdot \kappa} . \tag{21}$$

In this generality, these estimates are the best possible ones (see [23]). They show the importance of having well-conditioned splittings. Improved results for the multiplicative method can be found in, e.g., [32, 34, 1]. If the splitting is of multilevel type (16) then the iterations **(AS)** and **(MS)** can be interpreted as a V-cycle preconditioner and a V-cycle multigrid algorithm, respectively. For a comprehensive treatment of multigrid theory in the framework of subspace correction algorithms, see [1]. Additive multilevel preconditioning, especially based on multilevel frames and Riesz bases in Sobolev spaces, is emphasized in [22, 6, 7, 27]. Domain decomposition algorithms are treated in [3, 31, 33]. The concept of space splittings has also been applied to discretizations for elliptic systems (Stokes and Maxwell equations). Generalizations to nonsymmetric, indefinite, and unstructured problems have been attempted, with mixed success.

4 Multilevel schemes for the single layer potential equations

The *single layer potential equation*

$$Tf \equiv \frac{1}{4\pi} \int_\Gamma \frac{f(y)}{|x-y|} \, dy = g(x) , \tag{22}$$

is the prototype of an operator equation associated with a symmetric elliptic pseudodifferential operator $T : H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ of order -1. One concrete application are capacity calculations of electrically charged bodies in $\mathbf{R}^3$ where Γ is the surface of the body, $g(x) = 1$ on Γ, and $f(x)$ represents the charge density. The capacity itself is defined by

$$C = C(\Gamma) = \frac{1}{4\pi} \int_\Gamma f(y) \, dy \tag{23}$$

Similar problems arise for *interface problems* for second order elliptic boundary value problems when Neumann data have to be determined from Dirichlet data.

The numerical analysis of an integral equation such as (22) faces a number of difficulties:

- The solution theory "lives" in $H^{-1/2}(\Omega)$, the natural energy space for (22). Sobolev norms of negative order are not so well-investigated in connection with standard discretization schemes.
- For non-smooth bodies, e.g., if Γ contains corners and edges, the solutions exhibit strong singularities.
- The stiffness matrices A_N associated with any typical discretization space V_N (e.g., boundary element or spectral approximations) are dense matrices. Moreover, almost independently of the used discretization spaces and methods, the reliable numerical computation of the entries of the stiffness matrices represents a serious bottleneck. The singularity of the kernel $k(x,y) = 1/|x-y|$ along the 'diagonal' $x = y$ and the parametrization of the surface Γ represent additional challenges.
- The condition of A_N grows with dim V_N. Although this growth is moderate compared with second order elliptic equations, preconditioning needs to be considered.

These challenges have attracted many researchers. In particular, *hp-methods* and *wavelet methods* (combined with *matrix compression*) are under investigation. See [6, 20] for some references.

In Talk 2, we first report theoretical and numerical results [24] on *preconditioning low-order boundary element discretizations* for (22) using semi-orthogonal wavelet splittings. This material is closely related to Talk 1, and further illustrates the machinery of stable splittings in the Sobolev context. In particular, we derive the exact asymptotical properties of multilevel preconditioners based on Haar-type L_2-orthogonal bases in the case of piecewise constant elements.

The main part of this talk is devoted to our recent joint work with Griebel and Schiekofer [20, 28] on using *sparse grid spaces* to partly overcome the above-mentioned difficulties. This second part is restricted to the case of a *square screen* $\Gamma = [0,1]^2$ (and generalizes to Γ composed of a few tensor-product faces). In this case, sparse grid spaces can be constructed by looking at special sections of *tensor-product systems* obtained from univariate spline systems. E.g., if $\{H_{j,i}\}$ denotes the one-dimensional Haar system on $[0,1]$ (the *level index* j is determined by the requirement that $H_{j,i}$ is constant on dyadic intervals of length 2^{-j}) then

$$\hat{V}_J = \text{span}\{H_{\mathbf{j},\mathbf{i}} = H_{j_1,i_1} \otimes H_{j_2,i_2} \,:\, j_1 + j_2 \le J\}, \quad J \ge 0\,,$$

is the definition of standard sparse grid spaces for piecewise constant functions on $[0,1]^2$. If V_J denotes the standard *full grid spaces* (piecewise constant functions on a square grid of sidelength 2^{-J}) then $\hat{V}_J \subset V_J$ and, most importantly, dim $\hat{V}_J \approx J2^J << 2^{2J} = \dim V_J$. As is well-known for H^s-approximation with $s \ge 0$, under additional regularity assumptions (existence of certain higher order mixed derivatives) the use of $\hat{V}_J$ instead of V_J leads to good approximation rates with a small number of degrees of freedom. We make the approximation power of $\hat{V}_J$ precise for $s < 0$, and observe some reduced efficiency of the sparse grid approach for this case.

Our numerical experiments for Galerkin discretizations of (22) in the case $s = -1/2$, however, looked much more promising than predicted by the error analysis. We traced this phenomenon back to the favorable properties of tensor-product systems (as considered in the construction of sparse grid spaces) for the resolution of *edge singularities*. It is shown in [28] to which extent optimally constructed *adaptive sparse grid spaces* may be superior over traditional adaptive wavelet spaces in the case of (22). In particular, we can show that we can choose $\le N$ Haar functions $H_{\mathbf{j},\mathbf{i}}$ such that using them as ansatz functions in a Galerkin scheme for the capacity problem ($g(x) = 1$, $\Gamma = [0,1]^2$ in (22)) leads to capacity approximations C_N with an error rate of

$$|C - C_N| = \mathrm{O}(N^{-5/2})\,, \qquad N \to \infty\,.$$

Traditional adaptive wavelet schemes cannot reach rates better than $O(N^{-1})$ for the same problem. The practical results for capacity computations are very impressive if moderate accuracy (relative error $\approx 10^{-3}$) suffices (asymptotically, for very high resolution, hp-methods will be superior). This topic is related to investigations on nonlinear best N-term approximation [10], and to the problem of how to properly incorporate *anisotropic refinement*. It should be mentioned that most of our observations are not yet fully practical, and need further evaluation.

References

[1] J. H. Bramble, Multigrid Methods, Pitman Research Notes in Mathematical Sciences v. 294, Longman Sci.&Techn., Harlow, Essex, 1993.

[2] P. L. Butzer, K. Scherer, Approximationsprozesse und Interpolationsmethoden, Bibliogr. Institut, Mannheim, 1968.

[3] T. Chan, T. Mathew, Domain decomposition methods, Acta Numerica 94 (1994), 61–143.

[4] O. Christensen, Frames and the projection method. J. Appl. Comput. Harm. Anal. 1 (1993), 50–53.

[5] C. K. Chui, An Introduction to Wavelets, Academic Press, Boston, 1992.

[6] W. Dahmen, Wavelet and multiscale methods for operator equations, Acta Numerica 6 (1997), 55–228.

[7] Dahmen, W., A. Kurdila, P. Oswald (eds.), Multiscale Wavelet Methods for Partial Differential Equations, Academic Press, San Diego, 1997.

[8] W. Dahmen, R. Stevenson, Element-by-element construction of wavelets satisfying stability and moment conditions, IGPM-Report 145, RWTH Aachen, November 1997.

[9] I. Daubechies, Ten Lectures on Wavelets, CBMS-NSF Reg. Conf. Ser. Appl. Math. v. 61, SIAM, Philadelphia, 1992.

[10] R. DeVore, Nonlinear approxiamtion, Acta Numerica 7 (1998).

[11] M. Dryja, O. Widlund, Towards a unified theory of domain decomposition for elliptic problems. In: Proc. 3rd Int. Symp. on DDM for PDE (T. Chan, R. Glowinski, J. Periaux, O. Widlund, eds.), SIAM, Philadelphia, 1990.

[12] R. Duffin, A. Schaeffer, A class of nonharmonic Fourier series, TAMS 72 (1952), 341–366.

[13] H.-G. Feichtinger, T. Strohmer (eds.), Gabor Analysis and Algorithms: Theory and Applications, Birkhauser, Basel, 1997.

[14] K. Groechenig, Describing functions: atomic decompositions versus frames, Monatsh. Math. 112 (1992), 1–42.

[15] W. Hackbusch, Iterative Solution of Large Sparse Systems of Equations. Appl. Math. Sci. vol. 95, Springer, New York, 1994.

[16] C. Heil, D. F. Walnut, Continuous and discrete wavelet transform, SIAM Review 31 (1989), 628–666.

[17] M. Griebel, Multilevelverfahren als Iterationsverfahren über Erzeugendensystemen, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.

[18] M. Griebel and P. Oswald, Remarks on the abstract theory of additive 8 and multiplicative Schwarz methods, Numer. Math. 70 (1995), 163–180.

[19] M. Griebel and P. Oswald, Tensor-product-type subspace splittings and multilevel iterative methods for anisotropic problems, Adv. Comput. Math. 4 (1995), 171–206.

[20] M. Griebel, P. Oswald, T. Schiekofer, Sparse grids for boundary integral equations, Numer. Math. (in revision).

[21] R. Lorentz, P. Oswald: Criteria for hierarchical bases for Sobolev spaces. GMD-Report Nr. 1059, GMD, Sankt Augustin, March 1997 (submitted to ACHA).

[22] P. Oswald, Multilevel Finite Element Approximation: Theory and Application, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.

[23] P. Oswald, On the convergence rate of SOR: A worst case estimate. Computing 52 (1994), 245–255.

[24] P. Oswald, Multilevel norms for $H^{-1/2}$, Computing (accepted).

[25] P. Oswald, Multilevel splittings for interface problems, Proc. DD11, London, 1998.

[26] P. Oswald, Frames and space splittings in Hilbert spaces, Manuscript, September 1997.

[27] P. Oswald, Multilevel frames and Riesz bases in Sobolev spaces, Manuscript, June 1998.

[28] P. Oswald, Best N-term approximation with Haar functions, (in preparation).

[29] M. R. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, L. Raphael (eds.), Wavelets and Their Applications, Jones & Bartlett, Boston, 1992 (Chapter V).

[30] J. J. Benedetto, M. W. Frazier (eds.), Wavelets: Mathematics and Applications, CRC Press, Boca Raton, 1994.

[31] B. F. Smith, P. E. Bjorstad, W. D. Gropp, Domain Decomposition - Parallel Multilevel Methods for Elliptic Partial Differential Equations, Cambridge Univ. Press, Cambridge, 1996.

[32] J. Xu, Iterative methods by space decomposition and subspace correction, SIAM Review 34 (1992), 581–613.

[33] J. Xu, J. Zou, Some non-overlapping domain decomposition methods, SIAM Review (1998), (to appear).

[34] H. Yserentant, Old and new convergence proofs for multigrid methods, Acta Numerica 93, Cambr. Univ. Press, 1993, 285–326.

Element-by-element construction of wavelets satisfying stability and moment conditions

Rob Stevenson, University of Nijmegen

Multiscale- or wavelet bases can be viewed as improved hierarchical bases, known from finite element spaces, in the sense that they are *stable* in a range of Sobolev norms. As a consequence, stiffness matrices corresponding to elliptic problems with respect to these wavelet bases have uniformly bounded condition numbers not only for problems of second order, as is (almost) the case with the hierarchical basis (in 2D), but also for problems of lower or even negative orders, for example for various integral equations arising from the application of the boundary integral method.

In addition, unlike hierarchical basis functions, wavelets have *vanishing moments*, i.e., they are orthogonal to all polynomials of degree less than m, where m is the number of vanishing moments. A well-known disadvantage of the boundary integral method is that it gives rise to dense matrices. Yet, as a consequence of the vanishing moments, when wavelet bases are applied many elements in the matrix appear to be small. It has been shown that, dependent on the order of the equation and the order of approximation of the discrete space, when the number of vanishing moments is large enough, the stiffness matrix can be compressed to a sparse one, where the order of convergence of the resulting solution is retained. Since the compressed stiffness matrix is also well-conditioned, a method of optimal complexity is obtained for solving integral equations.

"Classical" wavelet spaces are spanned by the translates and dilates of one, or in more dimensions a few "mother wavelets". This means that applications are restricted to uniform meshes. Even the adaption of such wavelets to a bounded interval is a far from trivial task.

In this talk, we demonstrate a construction of wavelet bases of standard Lagrange finite element spaces on non-uniform meshes on n dimensional domains or manifolds. The wavelet bases are stable in the Sobolev spaces H^s for $|s| < \frac{3}{2}$ ($|s| \leq 1$ on Lipschitz' manifolds), and the wavelets can, in principal, be arranged to have any desired order of vanishing moments.

Based on the affine equivalence of the finite elements the construction of the wavelets consists of two parts: An implicit part involving some computations on a reference element which, for each type of finite element space, have

to be performed only once. In addition there is an explicit part which takes care of the necessary adaptations of the wavelets to the actual mesh. The only condition we need for this construction to work is that the refinements of initial elements are uniform.

We will show that the wavelet bases can be implemented efficiently.

References

[1] W. Dahmen and R.P. Stevenson. Element-by-element contruction of wavelets satisfying stability and moment conditions. Technical Report 9725, University of Nijmegen, November 1997. Submitted to SIAM J. Numer. Anal.

STIFF OSCILLATORY SYSTEMS WITH RANDOM INITIAL DATA

Andrew Stuart
Scientific Computing and
Computational Mathematics Program,
Durand 257,
Stanford University,
Stanford CA94305-4040, USA.

Introduction

In the field of computational molecular dynamics stiff oscillatory systems, with broad frequency spectra, often arise. It is hence of interest to develop a theory of the numerical analysis for such problems. In the area of stiff dissipative systems the understanding of numerical algorithms has been greatly enhanced by the study of a variety of simple model problems [1]; here we introduce, and then study numerical methods for, several model problems in stiff oscillatory systems.

We start by introducing two very simple model problems for stiff oscillatory systems. Both comprise a linear superposition of $N \gg 1$ harmonic oscillators used as a forcing term for a scalar ODE. In the first case the initial conditions are chosen so that the forcing term approximates a delta function as $N \to \infty$ and in the second case so that it approximates white noise. In both cases the fastest natural frequency of the oscillators is $\mathcal{O}(N)$.

The model problems are integrated numerically in the stiff regime where the time-step Δt satisfies

$$N\Delta t = \mathcal{O}(1). \tag{1}$$

The convergence of the algorithms is studied in this case in the limit

$$N \to \infty \quad \text{and} \quad \Delta t \to 0. \tag{2}$$

For the white noise problem both strong and weak convergence are considered. This work may be found in [3].

We then describe a model for a particle immersed in a heat bath. This simple mechanical model is prototypical of many models in statistical mechanics: the overall system has large dimension $2N+2$ but for $N \gg 1$ the projection of the solution onto a low dimensional subspace, of dimension 2 here, is governed by an equation of dimension 2 in which several parameters characterizing the overall statistics of the remaining $2N$ variables remain but no other details appear. Since the $2N$ variables play this simple role in the projected variables it is natural to ask whether it is necessary to resolve accurately those $2N$ variables if information in the projected space is all that is required. These ideas are closely related to those studied for the two simple model problems and we present numerical experiments to show that underresolved computations for the heat bath/particle model can still yield accurate simulations for certain projections of the solution. This work may be found in [4].

The Two Simple Model Problems

Consider the equations

$$\begin{gathered} \ddot{u}_j + j^2 u_j = 0, \\ u_j(0) = a_j, \quad \dot{u}_j(0) = 0, \quad j = 0, \ldots, N \end{gathered} \tag{3}$$

and

$$\begin{gathered} \dot{z}_N = f(Z_N) + H_N(t), \\ Z_N(0) = z_0 \end{gathered} \tag{4}$$

where

$$H_N(t) := \sum_{j=0}^{N} u_j(t). \tag{5}$$

We consider two choices for the $\{a_j\}_{j=0}^{N}$: the first is

$$a_0 = \frac{1}{2}, \quad a_j = 1,\ j \geq 1. \quad [MP1] \tag{6}$$

The second is

$$a_0 = \frac{1}{\sqrt{\pi}}\eta_0, \quad a_j = \sqrt{\frac{2}{\pi}}\eta_j,\ j \geq 1; \quad [MP2] \tag{7}$$

here the η_j are IID Gaussian random variables with mean 0 and variance 1.

Throughout the following we assume that $f \in C^\infty(\mathbb{R}^m, \mathbb{R}^m)$ satisfies the global bounds

$$\left\{ \begin{array}{rcl} \|f(x) - f(y)\| & \leq & L\|x - y\| \\ \|f(x)\| & \leq & K[1 + \|x\|] \end{array} \right\} \quad \forall x, y \in \mathbb{R}^m.$$

Formal calculations indicate that for [MP1], $0 \leq t \leq \pi$ and N large, Z_N should behave like z solving

$$\dot{z} = f(z), \quad z(0) = z_0 + \frac{\pi}{2}. \tag{8}$$

For [MP2] the analogous formal limit is the SDE

$$dz = f(z)dt + dW, \quad z(0) = z_0 \tag{9}$$

where W is a standard Brownian motion on $0 \leq t \leq \pi$. The following three results make this intuition precise.

Theorem 1 *Consider $Z_N(t)$ solving* [MP1] *and $z(t)$ solving* (8). *Then, for $T \in [0, \pi]$,*

$$\|z(\cdot) - Z_N(\cdot)\|^2_{L^2(0,T)} \leq \frac{C(T)}{N}.$$

Theorem 2 *Consider $Z_N(t)$ solving* [MP2] *and $z(t)$ solving* (9). *Then, for $T \in [0, \pi]$,*

$$\mathbb{E}\|z(\cdot) - Z_N(\cdot)\|^2_{L^\infty(0,T)} \leq \frac{C(T)}{N}.$$

It is often the case that weak convergence results can be obtained at faster rates than strong convergence and we now demonstrate this. We consider expectations of functions $g : \mathbb{R} \to \mathbb{R}$ whose Fourier transform $\hat{g}$ satisfies

Hypothesis H *There exists a real number $\beta > 1$ and a positive constant C_1 such that*

$$|\hat{g}(k)| \leq C_1(1 + |k|)^{-\beta} \qquad \forall k \in \mathbb{R}. \tag{10}$$

In the following Theorem we consider the case $f \equiv 0$. Thus z solving (9) is a pure Brownian motion. This allows relatively straightforward analysis using Fourier techniques; more complicated methods would be required to analyze the case of non-zero f.

Theorem 3 *Let $f(z) \equiv 0$ and let $g : \mathbb{R} \to \mathbb{R}$ satisfy Hypothesis H. Consider $Z_N(t)$ solving* [MP2] *and $z(t)$ solving* (9). *Then, for $T \in [0, \pi]$,*

$$\sup_{z_0 \in R} |\mathbb{E}g(z(T)) - \mathbb{E}g(z_N(T))| \leq \begin{cases} CN^{(1-\beta)/2}, & 1 < \beta < 3, \\ CN^{-1}\log(1+N), & \beta = 3, \\ CN^{-1}, & \beta > 3, \end{cases} \tag{11}$$

where $C = C(\beta, C_1)$ with β and C_1 as in Hypothesis H.

The Heat Bath Model

In this section we describe a simplified model for the statistical mechanics of a heat bath, taken from [2]. Consider the Hamiltonian for a single *distinguished particle* of unit mass moving in a potential V and attached by linear springs to N harmonic oscillators each with mass m_j and stiffness k_j :

$$H = \frac{1}{2}p^2 + V(q) + \sum_{j=1}^{N} \left\{ \frac{v_j^2}{2m_j} + \frac{k_j}{2}(u_j - q)^2 \right\}. \tag{12}$$

The Hamiltonian (12) gives rise to Hamilton's equations

$$\begin{aligned} \dot{p} &= -V'(q) + \textstyle\sum_{j=1}^{N} k_j(u_j - q), \\ \dot{q} &= p, \\ \dot{v}_j &= -k_j(u_j - q), \\ \dot{u}_j &= v_j/m_j. \end{aligned} \tag{13}$$

Here the u_j, v_j represent the heat bath and the variables p, q the particles which is in thermal contact with the bath.

Under certain natural conditions on the spring constants and masses, and for certain random initial data (from the Boltzmann distribution) it may be shown, by eliminating the heat-bath variables u_j, v_j, that q satisfies the equation

$$\ddot{q} + V'(q) + \int_0^t K_N(t-s)\dot{q}(s)ds = -K_N(t)q(0) + Z_N(t) \tag{14}$$

where

$$\begin{aligned} K_N(t) &= \gamma^2 \textstyle\sum_{j=1}^{N} \cos(jt), \\ Z_N(t) &= \frac{\gamma}{\sqrt{\beta}} \textstyle\sum_{j=1}^{N} \mu_j \cos(jt) \end{aligned} \tag{15}$$

where $\{\mu_j\}_{j=1}^{\infty}$ are IID random variables distributed as $\mathcal{N}(0,1)$.

Formally (and this can be made precise) taking the limit $N \to \infty$ yields the candidate limit problem

$$\begin{aligned} &\ddot{Q} + \tfrac{\gamma^2\pi}{2}\dot{Q} + V'(Q) - \tfrac{\gamma^2}{2}Q = \dot{W}, \\ &Q(0) = q_0, \quad \dot{Q}(0) = p_0 - \tfrac{\gamma^2\pi}{2}q_0. \end{aligned} \tag{16}$$

Here $\dot{W}$ (the limit of Z_N) is closely related to white noise so that a precise interpretation of this equation requires reformulation as an integral equation.

We can now ask questions analogous to those considered in the previous section. We solve the large system generated by the Hamiltonian (12) in the regime (1) and consider the limit (2). Numerical experiments will show that the theory developed for the simple model problems is instructive in understanding this more complicated model of a heat bath.

We study a parameterized family of numerical methods applied to the Hamiltonian system of dimension $2N + 2$. These methods are constructed to be energy conserving for the homogeneous part of the heat bath. We fix the product of the time-step and largest natural frequency at $\mathcal{O}(1)$. The dimension of the problem is then increased ($N \to \infty$) and the computed solutions for the distinguished particle are compared with the exact motion given by the SDE ($N = \infty$). In this set-up the fastest scales are not accurately resolved and it is of interest to ask whether the (macroscopic) motion of the distinguished particle is, nonetheless, accurately resolved. We show formally that, in the underresolved regime, the computed motion of the distinguished particle approximately satisfies an SDE whose coefficients depend on the parameters defining the method. For certain combinations of parameters this SDE agrees with the true SDE governing the motion of the distinguished particle and these are the methods which compute the correct limiting behaviour as $N \to \infty$. For other combinations of parameters the computed SDE limit has different damping (possibly negative) and different initial conditions from the true SDE limit. We give numerical experiments which support these results, together with experiments demonstrating the fact that the backward Euler method reproduces the correct behaviour under the aforementioned limit process with $N \to \infty$ and $\Delta t \to 0$. Note that the backward Euler method allows far larger time-steps to be taken than for the other methods considered here, and this is the motivation for its study.

References

[1] **K. Dekker and J.G. Verwer**, STABILITY OF RUNGE-KUTTA METHODS FOR STIFF NONLINEAR EQUATIONS, North Holland, Amsterdam, 1980.

[2] **G.W. Ford and M. Kac**, ON THE QUANTUM LANGEVIN EQUATION. J. Stat. Phys. **46**(1987), 803–810.

[3] **B. Cano, A.M. Stuart, E. Suli and J.O. Warren** STIFF OSCILLATORY SYSTEMS, DELTA JUMPS AND WHITE NOISE. In preparation.

[4] **A.M. Stuart and J.O. Warren**, ANALYSIS AND EXPERIMENTS FOR A COMPUTATIONAL MODEL OF A HEAT BATH. Submitted to SISC.

PERTURBATION THEORY FOR ERGODIC MARKOV CHAINS

Andrew Stuart
Scientific Computing and
Computational Mathematics Program,
Durand 257,
Stanford University,
Stanford CA94305-4040, USA.

Introduction

We consider approximation of the SDE in $\mathbb{R}^m$

$$du = f(u)dt + \sigma(u)dW, \quad u(0) = x.$$

Here $f : \mathbb{R}^m \to \mathbb{R}^m$, W is a $d-$dimensional Brownian motion and, for each $u \in \mathbb{R}^m$, $\sigma(u) : \mathbb{R}^d \to \mathbb{R}^m$. Our aim is to understand the approximation of such equations over long time intervals. To be definite consider the approximation $U^n \approx u(t^n), t^n = n\Delta t$ with

$$U^{n+1} = U^n + \Delta t f(U^n) + \sigma(U^n)\Delta W^n, \quad U^0 = x$$

where ΔW^n is the ($d-$dimensional Gaussian) increment $W(t^{n+1}) - W(t^n)$. For background materials on numerics for SDEs see [1]. We refer to the original SDE as (P) and its approximation as (PD).

The Deterministic Case $\sigma \equiv 0$

Even in the deterministic case ($\sigma = 0$) understanding the effect of approximation on long-time behaviour is a hard problem and we start be surveying briefly some of the known results in that case, concerning approximation of long-time properties – see [4] and [5] for details. The key point is that trajectories of the equation (P) and its approximation, in general, diverge:

$$\|u(t^n) - U^n\| \leq c_1 e^{c_2 T} \Delta t^r, \quad 0 \leq n\Delta t \leq T.$$

Thus we must look at other objects if we wish to understand the sense in which (P) and (PD) are close over long time-intervals. Invariant sets $\mathcal{I}$ satisfy

$$S(\mathcal{I}, t) \equiv \mathcal{I}$$

and play a fundamental role in the dynamics of (P) – equilibria, periodic solutions, invariant manifolds and general attractors are all invariant sets. Thus rather than looking at trajectories attention is focussed on the objects around which the dynamics are organized.

Many strong results can be proved about the existence and closeness of approximating invariant sets $\mathcal{I}_{\Delta t}$ in (PD) when some form of local hyperbolicity is present in for $\mathcal{I}$. These results apply to, for example, periodic solutions, quasi-periodic solutions, invariant manifolds etc. A natural non-hyperbolic object to study is an attractor $\mathcal{A}$: a compact invariant set which attracts a neighbourhood of itself. If (P) has an attractor $\mathcal{A}$ then (PD) has an attractor $\mathcal{A}_{\Delta t}$ and, for any $\epsilon > 0$, there is $\Delta t_c(\epsilon)$ such that

$$\mathcal{A}_{\Delta t} \subseteq \mathcal{N}(\mathcal{A}, \epsilon) \quad \forall \Delta t < \Delta t_c(\epsilon).$$

This *upper semicontinuity* result is the best that can be said, in general, and obstruction to further generality is caused by a lack of hyperbolicity. The techniques are similar to those encountered for parametric perturbations of vector fields, but care is required, especially for PDEs, to ensure that error estimates in appropriate spaces are used.

To prove *lower semicontinuity*:

$$\mathcal{A} \subseteq \mathcal{N}(\mathcal{A}_{\Delta t}, \epsilon) \quad \forall \Delta t < \Delta t_c(\epsilon).$$

it is necessary to make further assumptions about the attractor $\mathcal{A}$. The simplest is that $\mathcal{A}$ comprises the closure of the union of unstable manifolds of equilibria.

Ergodicity – Deterministic Case

Many dynamical systems are thought to be *ergodic* in the sense that, as $T \to \infty$,

$$\frac{1}{T} \int_0^T g(S(x,t))dt \to \int_X g(x)\mu(dx)$$

where X is some invariant set of (P) (e.g. $\mathcal{A}$, an attractor) and μ is a measure supported on X. In such an instance it is natural to revisit the question of

convergence of trajectories and study whether

$$E := \|\frac{1}{T}\int_0^T g(S(x,t))dt - \frac{1}{N}\sum_{n=0}^{N-1} g(S^n_{\Delta t}(x))\|$$

is small for $T = N\Delta t \gg 1$; note that this is certainly not implied by the standard finite time convergence result. However, intuitively, time-averaging should help to prevent accumulation of errors.

Proving results of this type appears to be extremely hard in the deterministic case, even though there is strong numerical evidence to suggest that positive results are likely in a wide range of situations. The only result currently available in this area is due to Reich [2] who works in the context of certain hyperbolic flows governed by Hamiltonian systems. Reich uses the ideas of shadowing and backward error analysis to show that symplectic methods approximate time averages well for *exponentially long* (in Δt) periods of time:

$$E = \mathcal{O}(\Delta t^p), \quad 1 \ll N\Delta t \leq \mathcal{O}(e^{c/\Delta t})$$

– far better than the logarithmic time (in Δt) implied by standard error estimates.

Ergodicity – Stochastic Case

It turns out that randomness actually makes the study of ergodicity easier. For problem (P) the equation governing propagation of probability densities (given random initial data and/or random stochastic forcing through σ) is of the form

$$\frac{\partial p}{\partial t} = \mathcal{L}^* p$$

and for the SDE $\mathcal{L}$ has some parabolic features whilst in the deterministic case $\sigma \equiv 0$ it is purely hyperbolic. The operator $\mathcal{L}$ is known as the generator of the stochastic process.

Our starting point is a generalization of the classical estimate for errors in the numerical approximation of ODEs: we assume that

$$\mathbb{E}\|u(t^n) - U^n\| \leq c_1 e^{c_2 T}\Delta t^r.$$

The question is whether we can go from this assumption to deduce long-time approximation properties for time-averages. In the random case considerably

more can be said concerning this question for ergodic problems, than in the deterministic case.

If we define

$$B(u) = \sigma(u)\sigma(u)^T$$

then the SDE has a uniformly parabolic generator $\mathcal{L}$ if

$$\exists c > 0 : \xi^T B(u)\xi \geq c\|\xi\|^2 \forall \xi, u \in \mathbb{R}^m;$$

it is shown by Talay [6] that in this case, a.s.,

$$E = \mathcal{O}(\Delta t^r),$$

for all $T = N\Delta t \gg 1$. For SDEs with non-uniformly parabolic generators, but for which exponential ergodicity holds, the weaker result that, a.s.,

$$E = \mathcal{O}(\Delta t^{\gamma r}),$$

for all $T = N\Delta t \gg 1$ and some $\gamma \in (0, 1)$. This is proved in [3].

References

[1] **P. Kloeden and E. Platen**, NUMERICAL SOLUTION OF STOCHASTIC DIFFERENTIAL EQUATIONS. Springer, 1994.

[2] **S. Reich**, *Backward error analysis* $\cdots$, PREPRINT, 1998.

[3] **T. Shardlow and A.M. Stuart**, *A perturbation theory for ergodic properties of Markov chains.* To appear SIAM J NUM. ANAL., 1998.

[4] **A.M. Stuart and A.R. Humphries**, DYNAMICAL SYSTEMS AND NUMERICAL ANALYSIS. Cambridge University press, 1996.

[5] **A.M. Stuart**, *Stability and convergence in the numerical approximation of dynamical systems.* Appears in STATE OF THE ART IN NUMERICAL ANALYSIS, eds; A. Iserles and G.A. Watson. Oxford University Press, 1997.

[6] **D. Talay**, *Second-order discretization schemes for stochastic differential systems for the computation of the invariant law.* STOCHASTICS AND STOCHASTIC REPORTS **29**(1990), 13–36.

M. Zennaro
Dipartimento di Scienze Matematiche
Università di Trieste
34100 Trieste (Italy)

1 FIRST LECTURE: An introduction to the numerical solution of delay differential equations

A large number of real life mathematical models are based on initial value problems (IVPs) for ordinary differential equations (ODEs) of the type

$$\begin{cases} y'(t) = f(t, y(t)), & t \geq t_0, \\ y(t_0) = y_0, \end{cases} \tag{1.1}$$

where the function $y(t)$ represents some physical quantity which evolves in time.

However, in order to make the model more consistent with the real phenomenon, it sometimes is necessary to modify the right-hand side of (1.1) to include the dependence of the derivative y' also on y computed at some past value $t - \tau$. According to the complexity of the phenomenon, the *delay* τ, which always is nonnegative, may be just a constant (*constant delay*), or a function of t (*variable delay*), or even a function of t and y itself (*state dependent delay*). In any case, equation (1.1) modifies to

$$\begin{cases} y'(t) = f(t, y(t), y(t-\tau)), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \tag{1.2}$$

which is called *delay differential equation* (DDE).

In more general models, the derivative y' may depend on y and y' itself at some past value $t - \tau$. In this case (1.1) changes into

$$\begin{cases} y'(t) = f(t, y(t), y(t-\tau), y'(t-\tau)), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \tag{1.3}$$

where the function $\phi(t)$ is supposed to be at least C^1-continuous. Equation (1.3) is called a *delay differential equation* of *neutral* type (NDDE).

A first difference between equations (1.1) and (1.2) - (1.3) is that the solution of the latter ones is determined by an initial function $\phi(t)$ rather than by a

simple initial value y_0, as happens for the former. As a consequence, even if the functions f, τ and ϕ in (1.2) - (1.3) are C^∞-continuous, in general the solution $y(t)$ is not smoothly linked to the initial function $\phi(t)$ at the point t_0, where only C^0-continuity can be assured. Such discontinuity is spread forward along the integration interval and a set of ***discontinuity points*** is generated, whose location is determined by the delayed argument $t - \tau$.

In these lectures we assume that the delay τ is either constant or variable, but not state dependent. Moreover, if it is variable, the following properties are assumed to be satisfied:

(H_1) there exists a constant $\tau_0 > 0$ such that $\tau(t) \geq \tau_0$ for all $t \geq t_0$;

(H_2) the delayed argument $t - \tau(t)$ is a strictly increasing function for all $t \geq t_0$.

Under these hypotheses, the discontinuity points are generated inductively by the recursion

$$\xi_k - \tau(\xi_k) = \xi_{k-1}, \quad k \geq 1, \tag{1.4}$$

where $\xi_0 = t_0$, and an increasing sequence $\{\xi_k\}_{k\geq 0}$ is determined which can actually be computed a priori by using (1.4). In this way, a sequence of intervals $[\xi_{k-1}, \xi_k]$ is also defined. Moreover, each pair of consecutive discontinuity points satisfies $\xi_k - \xi_{k-1} \geq \tau_0$.

The hypotheses (H_1) and (H_2) yield existence and uniqueness of the solution quite easily. In fact, they can be proved just by using induction on the intervals $[\xi_{k-1}, \xi_k]$ and the well known existence and uniqueness theorem for ODEs (1.1) under the hypothesis of uniform Lipschitz continuity of the right-hand side.

The discontinuity points ξ_k are sometimes called ***primary discontinuities***. If the functions f, τ and ϕ in (1.2) - (1.3) have some discontinuities with respect to t in some of their derivatives, then such discontinuities are also propagated by the delayed argument $t - \tau$ following the rule (1.4) and are called ***secondary discontinuities***. However, in order to simplify the discussion, we assume that all the functions in (1.2) are C^∞-continuous. Therefore, in the interior of each interval $[\xi_{k-1}, \xi_k]$ the solution y is C^∞-continuous as well, and no secondary discontinuities are present.

Moreover, it can easily be seen that, at each discontinuity point ξ_k, the solution y of the DDE (1.2) is at least k times continuously differentiable. In other words, there is a smoothing of the solution y at the discontinuity points ξ_k as the index k increases.

The situation is different for the the solution y of the NDDE (1.3), since the smoothing, in general, does not take place, leaving it only C^0-continuous at all discontinuity points ξ_k.

The first approaches to the numerical solution of DDEs and NDDEs go back to the early sixties and were characterized by the straight application of well known formulae for ODEs, essentially ***linear multistep*** (LM) ***methods***. In particular, a set $\Delta = \{t_0, t_1, \ldots, t_n, \ldots\}$ of mesh points was assumed to exist

such that, for all $t_n \in \Delta$, either $t_n - \tau(t_n) < t_0$ or $t_n - \tau(t_n) \in \Delta$. Once we have such a mesh available, any discrete method making use of nodal points only can be implemented directly. For example, the *forward Euler method* for equation (1.2) looks like

$$y_{n+1} = y_n + h_{n+1} f(t_n, y_n, y_q),$$

whereas, for equation (1.3), it looks like

$$\begin{aligned} y_{n+1} &= y_n + h_{n+1} f(t_n, y_n, y_q, y'_q), \\ y'_n &= f(t_n, y_n, y_q, y'_q), \end{aligned}$$

for some integer $q < n$. This approach entails a severe constraint on the mesh which, in some cases, makes the method impracticable.

Later it was proposed to free the mesh selection from the delay, and to use extranodal points for the approximation of the delayed term $y(t - \tau)$.

For the numerical solution of the DDE (1.2), the most elegant approach aimed to avoid the need of interpolation was proposed by Bellman. It was first developed for constant delay and, successively, for variable delays subject to the hypotheses (H_1) and (H_2). Although Bellman's method is probably not the most convenient approach for solving DDEs numerically, it is certainly attractive because it allows the use of variable stepsize without any interpolation.

Nowadays, the most common approach for solving (1.2) - (1.3) is to proceed step-by-step across a mesh $\Delta = \{t_0, t_1, \ldots, t_n, \ldots\}$ as follows. Once a continuous approximation $\eta(t)$ is obtained for $t \leq t_n$, the $(n+1)$-st step consists in solving numerically, by means of a *continuous numerical method*, the equation

$$\begin{cases} w'(t) = f(t, w(t), x(t - \tau(t))), & t_n \leq t \leq t_{n+1}, \\ w(t_n) = \eta(t_n), \end{cases} \tag{1.5}$$

(in case of (1.2)) or

$$\begin{cases} w'(t) = f(t, w(t), x(t - \tau(t))), z(t - \tau(t))), & t_n \leq t \leq t_{n+1}, \\ w(t_n) = \eta(t_n), \end{cases} \tag{1.6}$$

(in case of (1.3)), where

$$x(s) = \begin{cases} \phi(s) & \text{for} \quad s \leq t_0, \\ \eta(s) & \text{for} \quad t_0 \leq s \leq t_n, \\ w(s) & \text{for} \quad t_n \leq s \leq t_{n+1}, \end{cases}$$

and $z(s)$ is the derivative of $x(s)$ or any other approximation of $y'(s)$ such as, for example, $z(s) = f(s, \eta(s), x(s - \tau(s))), z(s - \tau(s)))$.

Observe that, for $t - \tau(t) \leq t_n$, (1.5) - (1.6) reduce to an ODE. On the contrary, when $t - \tau(t) > t_n$ for some $t \in [t_n, t_{n+1}]$, they are true functional differential equations, requiring a more complicate approach for their numerical treatment. However, under hypotheses (H_1) and (H_2) it is always possible to

select the stepsize in order that this case does not occur. For example, it is sufficient to choose a stepsize $\leq \tau_0$.

Once the equations (1.5) - (1.6) are solved for $w(t)$, the approximation y_{n+1} is set equal to the approximate solution w_{n+1} of $w(t_{n+1})$ and the continuous extension $\eta(t)$ is prolonged up to t_{n+1}.

In these lectures we address ourselves to the analysis of continuous Runge-Kutta (CRK) methods applied to (1.2) - (1.3).

We recall that, given a mesh $\Delta = \{t_0, t_1 \ldots, t_n, \ldots\}$, the CRK method for the solution of the ODE (1.1) is defined as follows:

$$Y^i_{n+1} = y_n + h_{n+1} \sum_{j=1}^{s} a_{ij} f(t^j_{n+1}, Y^j_{n+1}), \quad i = 1, \ldots, s, \tag{1.7}$$

$$y_{n+1} = y_n + h_{n+1} \sum_{i=1}^{s} b_i f(t^i_{n+1}, Y^i_{n+1}), \tag{1.8}$$

where $t^i_{n+1} := t_n + c_i h_{n+1}, c_i := \sum_{j=1}^{s} a_{ij}, i = 1, \ldots, s$, $h_{n+1} := t_{n+1} - t_n$ and s is the number of *stages*. The b_i's are called *weights* of the quadrature formula (1.8) and the c_i's are called *abscissae*. For most of common methods and, in any case, for the methods considered in these lectures, the abscissae belong to $[0, 1]$.

The continuous extension $\eta(t)$ is defined, in each subinterval of the mesh Δ, by a one-step continuous quadrature rule of the form

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{s} b_i(\theta) f(t^i_{n+1}, Y^i_{n+1}), \quad 0 \leq \theta \leq 1, \tag{1.9}$$

where the $b_i(\theta)$'s are polynomials of suitable degree $\leq d$ satisfying the continuity condition

$$b_i(0) = 0 \quad \text{and} \quad b_i(1) = b_i, \quad i = 1, \ldots, s.$$

When applied to the DDEs (1.2), the CRK method assumes the following form:

$$Y^i_{n+1} = y_n + h_{n+1} \sum_{j=1}^{s} a_{ij} f(t^j_{n+1}, Y^j_{n+1}, \eta(t^j_{n+1} - \tau)), \quad i = 1, \ldots, s, \tag{1.10}$$

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{s} b_i(\theta) f(t^i_{n+1}, Y^i_{n+1}, \eta(t^i_{n+1} - \tau)), \quad 0 \leq \theta \leq 1, \tag{1.11}$$

where the delay τ has to be evaluated at the relevant points t^j_{n+1}. Of course, for $t \leq t_0$ we define $\eta(t) = \phi(t)$.

Analogously, for the NDDE (1.3) the method looks like

$$Y_{n+1}^i = y_n + h_{n+1} \sum_{j=1}^{s} a_{ij} f(t_{n+1}^j, Y_{n+1}^j, \eta(t_{n+1}^j - \tau), \psi(t_{n+1}^j - \tau)), \quad i = 1, \ldots, s,$$

$$\eta(t_n + \theta h_{n+1}) = y_n + h_{n+1} \sum_{i=1}^{s} b_i(\theta) f(t_{n+1}^i, Y_{n+1}^i, \eta(t_{n+1}^i - \tau), \psi(t_{n+1}^j - \tau)), \quad 0 \le \theta \le 1,$$

where $\psi(t)$ is either the derivative $\eta'(t)$ of the continuous numerical solution or another continuous approximation to $y'(t)$.

We discuss the effect of the discontinuity points $\{\xi_k\}$ on the order of convergence of the methods and discuss how to handle them in order to prevent the loss of accuracy.

We give a general theorem about the global order p' of convergence of CRK methods for (1.2) - (1.3) which relates it to the global order p of the discrete RK method (applied to ODEs) and to order of uniform convergence of the continuous approximations $\eta(t)$ and $\psi(t)$.

2 SECOND LECTURE: Some stability problems for delay differential equation solvers

For applications, sometimes it may be interesting to consider some mathematical models based on delay differential equations (DDEs), possibly of neutral type (NDDEs), whose solutions show a stable asymptotic behaviour. Clearly, in such situations, also the numerical methods used for the approximate solution of the model equation should preserve the same qualitative characteristic.

In this lecture we present some of the simplest linear stable problems and discuss the main tools used to test the stability properties of numerical methods.

Linear problems may be divided into two main classes, either of which requires a different tool for the analysis of stability properties: *autonomous* and *nonautonomous* problems. This dichotomy holds both for the analytic problem and for the numerical method.

As it will be illustrated, autonomous problems may be treated by analyzing the *characteristic equation*, whereas the treatment of nonautonomous problems requires to pass through contractivity. For this reason, it is often possible to find sharper results for the former class of problems.

Indeed, stability results for nonautonomous problems can be obtained also by using suitable Lyapunov functionals, but we do not consider this technique in our lecture.

The simplest autonomous test problem for DDEs is represented by the following scalar equation with constant delay:

$$\begin{cases} y'(t) = \lambda y(t) + \mu y(t - \tau), & t \ge t_0, \\ y(t) = \phi(t), & t \le t_0, \end{cases} \tag{2.1}$$

where λ and μ are complex parameters.

It is known that the condition

$$\Re(\lambda) + |\mu| < 0, \quad t \geq t_0, \tag{2.2}$$

implies

$$\lim_{t \to +\infty} y(t) = 0. \tag{2.3}$$

The stability analysis may be done directly by studying the roots of the *characteristic equation*

$$\zeta - \lambda - \mu e^{-\tau\zeta} = 0. \tag{2.4}$$

Such an equation has infinitely many solutions ζ_i, each of which with a certain multiplicity m_i, that do not accumulate anywhere in the complex plane.

It is known that the solution to (2.1) has an expansion of the form

$$y(t) = \sum_{i=1}^{\infty} \sum_{n_i=0}^{m_i-1} \alpha_{in_i} t^{n_i} e^{\zeta_i t}, \tag{2.5}$$

where the coefficients α_{in_i} are determined by the initial function $\phi(t)$. In view of the representation (2.5), it is easy to understand that a necessary and sufficient condition for the asymptotic stability of (2.1) is that all the roots ζ_i of (2.4) be such that $\Re(\zeta_i) < 0$. It is easy to verify that such a condition is guaranteed if (2.2) holds.

The successive step is to consider linear autonomous systems of the form

$$\begin{cases} y'(t) = Ly(t) + My(t-\tau), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \tag{2.6}$$

where L and M are constant complex $m \times m$-matrices.

The characteristic equation is

$$\det(\zeta I - L - e^{-\tau\zeta} M) = 0. \tag{2.7}$$

The discussion of the sign of the real part of the roots of (2.7) is more difficult than for (2.4). However, it can be seen that, if the matrices L and M are simultaneously diagonalizable, then a suitable change of variable involving the common eigenvectors leads to a decoupled system of m independent scalar equations of the form (2.1). Therefore the asymptotic stability of the system (2.6) is, in this case, completely described by the eigenvalues of the two matrices. More complicated is the case when the eigenspaces of L and M are different. It has been proved that a sufficient condition for the asymptotic stability of (2.6) is:

$$\Re(\lambda) < 0 \ \forall \text{ eigenvalues } \lambda \text{ of } L \quad \text{and} \quad \sup_{\Re(\xi)=0} \rho[(\xi I - L)^{-1} M] < 1, \tag{2.8}$$

where $\rho[\cdot]$ denotes the spectral radius of a matrix.

It is worth remarking that the sufficient conditions (2.2) and (2.8) are not necessary for asymptotic stability of (2.1) and (2.4), respectively, if the constant delay τ has a fixed value. On the contrary, they become (almost) necessary when we want to assure stability for all possible positive constant delays.

The method of the characteristic equation can be applied also to the analysis of neutral test equations. Similar results have, in fact, been proved for the following linear autonomous problems:

$$\begin{cases} y'(t) = \lambda y(t) + \mu y(t-\tau) + \nu y'(t-\tau), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \tag{2.9}$$

where λ, μ and ν are complex parameters, and

$$\begin{cases} y'(t) = Ly(t) + My(t-\tau) + Ny'(t-\tau), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \tag{2.10}$$

where L, M and N are constant complex $m \times m$-matrices. In particular, it has been proved that the condition

$$|\lambda\bar{\nu} - \bar{\mu}| + |\lambda\nu + \mu| < -2\Re(\lambda)$$

implies the asymptotic stability for all the solutions of (2.9).

Among the simplest nonautonomous problems are the following scalar linear equations:

$$\begin{cases} y'(t) = \lambda y(t) + \mu(t) y(t-\tau(t)), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \tag{2.11}$$

and

$$\begin{cases} y'(t) = \lambda y(t) + \mu(t) y(t-\tau(t)) + \nu(t) y'(t-\tau(t)), & t \geq t_0, \\ y(t) = \phi(t), & t \leq t_0, \end{cases} \tag{2.12}$$

where λ is a complex number with $\Re(\lambda) < 0$, $\mu(t)$ and $\nu(t)$ are continuous complex functions and the delay $\tau(t)$ satisfies the hypotheses (H_1) and (H_2).

By using induction on the sequence of intervals $[\xi_{k-1}, \xi_k]$, for the DDE (2.11) it is possible to show that the condition

$$\Re(\lambda) + |\mu(t)| \leq 0, \quad t \geq t_0,$$

implies the contractivity property (with respect to the initial data)

$$|y(t)| \leq \max_{x \leq t_0} |\phi(x)|, \quad t \geq t_0, \tag{2.13}$$

and that the stronger condition

$$R \cdot \Re(\lambda) + |\mu(t)| \leq 0, \quad t \geq t_0,$$

for some $R < 1$, imply the asymptotic stability property (2.3).

As for the NDDE (2.12), it is possible to show that the condition

$$\Re(\lambda)(1 - |\nu(t)|) + |\nu(t)\lambda + \mu(t)| \leq 0, \quad t \geq t_0,$$

implies the contractivity property

$$|y(t)| \leq \max\{|\phi(t_0)|; \max_{t_0 \leq x \leq \xi_1} \frac{|\mu(x)\phi(x - \tau(x)) + \nu(x)\phi'(x - \tau(x))|}{-\Re(\lambda)}\}, \quad t \geq t_0, \tag{2.14}$$

and that the stronger conditions

$$|\nu(t)| \leq \nu_0 < 1 \quad \text{and} \quad R \cdot \Re(\lambda)(1 - |\nu(t)|) + |\nu(t)\lambda + \mu(t)| \leq 0 \quad t \geq t_0,$$

for some $R < 1$, again imply the asymptotic stability property (2.3).

For both test equations (2.11) and (2.12), the proof of the above results is based on the fact that, for the scalar ODE with forcing term

$$\begin{cases} y'(t) = \lambda y(t) + g(t), & t \geq t_0, \\ y(t_0) = y_0, \end{cases} \tag{2.15}$$

where $g(t)$ is a continuous function and $\Re(\lambda) < 0$, it holds that

$$|y(t)| \leq e^{\Re(\lambda)(t-t_0)}|y_0| + (1 - e^{\Re(\lambda)(t-t_0)}) \max_{t_0 \leq x \leq t} \frac{|g(x)|}{-\Re(\lambda)}, \quad t \geq t_0.$$

In fact, the DDE (2.11) can be equivalently rewritten in the form (2.15) with $y_0 = \phi(t_0)$ and

$$g(t) = \begin{cases} \mu(t)\phi(t - \tau(t)), & t_0 \leq t \leq \xi_1, \\ \mu(t)y(t - \tau(t)), & t \geq \xi_1, \end{cases} \tag{2.16}$$

whereas the NDDE (2.12) can be equivalently rewritten in the form (2.15) with $y_0 = \phi(t_0)$ and

$$g(t) = \begin{cases} \mu(t)\phi(t - \tau(t)) + \nu(t)\phi'(t - \tau(t)), & t_0 \leq t \leq \xi_1, \\ (\nu(t)\lambda + \mu(t))y(t - \tau(t)) + \nu(t)g(t - \tau(t)), & t \geq \xi_1. \end{cases} \tag{2.17}$$

As far as numerical methods are concerned, considering the test equation (2.1) leads to the following generalization of the concept of *A-stability*.

Definition 2.1 *The* P-stability region *of a numerical step-by-step method for DDEs is the set S_P of the pairs of complex numbers (α, β), $\alpha = h\lambda$, $\beta = h\mu$, such that the discrete numerical solution $\{y_n\}_{n \geq 0}$ of (2.1) obtained with the constant stepsize h under the constraint*

$$h = \tau/m, \tag{2.18}$$

where m is a positive integer, satisfies

$$\lim_{n\to\infty} y_n = 0$$

for all constant delays τ and all initial functions $\phi(t)$. In particular, a numerical step-by-step method for DDEs is P-stable *if*

$$S_P \supseteq \{(\alpha,\beta) \in \mathbb{C}^2 \mid \Re(\alpha) + |\beta| < 0\}.$$

In other words, a numerical method for DDEs is P-stable if it preserves the asymptotic stability properties of the solution $y(t)$ of (2.1), under the constraint (2.18) on the stepsize.

Removing the constraint (2.18) leads to the similar but stronger concepts of *GP-stability* region and *GP-stable* method.

Analogous definitions, namely *NP-stability* and *GNP-stability*, are given with respect to the neutral test equation (2.9).

In this lecture we present some of the main results regarding the CRK methods for DDEs and NDDEs. We shall see that the key point is always the analysis of the solutions of vector difference equations (with constant coefficients) of arbitrarily high order and, in turn, of the roots of their characteristic equations. We shall also see that a suitable choice of the continuous extension $\eta(t)$ is sufficient to assure that any A-stable RK method (for ODEs) is also P-stable and NP-stable and, by using a special kind of multistep interpolation procedure, even GP-stable and GNP-stable.

Then we shall briefly illustrate how these results extend to the test systems (2.6) and (2.10).

The last part of the lecture is devoted to the analysis of the contractivity and asymptotic stability properties of CRK methods applied to the problems (2.15) - (2.16) and (2.15) - (2.17), which are equivalent to the test DDE (2.11) and NDDE (2.12), respectively.

This time, the nonautonomous character of the resulting difference equations makes characteristic equations useless for studying stability. Therefore, like for the analysis of the true solutions, induction on the sequence of intervals $[\xi_{k-1}, \xi_k]$ must be employed, together with the following particular contractivity property of the numerical method.

Definition 2.2 *The continuous RK method (1.7) - (1.8) - (1.9) is* A_f-stable *if the continuous numerical solution $\eta(t)$ of (2.15) satisfies*

$$|\eta(t_n + \theta h_{n+1})| \le \max\{|y_n|, \max_{1\le i\le s} \frac{|g(t_{n+1}^i)|}{-\Re(\lambda)}\}, \quad 0 \le \theta \le 1,$$

whenever $\Re(\lambda) < 0$ and for any mesh Δ.

We shall see that A_f-stability assures that the CRK method preserves the contractivity properties (2.13) and (2.14) and, also, asymptotic stability.

REFERENCES

For most of the results reported in these lectures we refer to the bibliography contained in

M. Zennaro:*Delay differential equations: theory and numerics*, in Theory and Numerics of ordinary and partial differential equations, M. Ainsworth, J. Levesley, W.A. Light, and M. Marletta eds, Clarendon Press, Oxford, 1995

List of participants Woudschoten Conference 23–25 September 1998

1.	Werner	Aernouts	KU Leuven	Leuven
2.	Martijn	Anthonissen	TU Eindhoven	Eindhoven
3.	J.-P.	Antoine	Univ. Cath. de Louvain	Louvain-la-Neuve
4.	Guido	vanden Berghe	Univ. Gent	Gent
5.	Patrick	Berkvens	CWI	Amsterdam
6.	Wim	Bomhof	Univ Utrecht	Utrecht
7.	Natalia	Borovykh	RU Leiden	Leiden
8.	Mike	Botchev	CWI	Amsterdam
9.	Kevin	Burrage	Univ. of Queensland	St. Lucia, Brisbane
10.	J.J.G.	Buschgens	TU Eindhoven	Eindhoven
11.	W.	Dahmen	RWTH Aachen	Aachen
12.	Kees	Dekker	TU Delft	Delft
13.	Jos	van Dorsselaer	CWI	Amsterdam
14.	Wienand	Drenth	TU Eindhoven	Eindhoven
15.	Koen	Engelborghs	KU Leuven	Leuven
16.	Jan	van Eijkeren	RIVM	Bilthoven
17.	Jason	Frank	TU Delft	Delft
18.	Arjan	Frijns	TU Eindhoven	Eindhoven
19.	Jaap G.	Fijnvandraat	Philips ED-T/AS	Eindhoven
20.	Menno	Genseberger	CWI	Amsterdam
21.	Jaap	van de Griend	RU Leiden	Leiden
22.	Pieter	de Groen	VU Brussel	Brussel
23.	Edwin	Havik	KdV Inst-UvA	Amsterdam
24.	Tanja	van Hecke	Univ. Gent	Gent
25.	Piet	Hemker	CWI	Amsterdam
26.	Guido	van den Heuvel	RU Leiden	Leiden
27.	Michiel	Hochstenbach	Univ Utrecht	Utrecht
28.	Bas	van 't Hof	TU Eindhoven	Eindhoven
29.	Walter	Hoffmann	UvA	Amsterdam
30.	Karel	in 't Hout	RU Leiden	Leiden
31.	P.J.	van der Houwen	CWI	Amsterdam
32.	Willem	Hundsdorfer	CWI	Amsterdam
33.	Jos K.M.	Jansen	TU Eindhoven	Eindhoven
34.	Rik	Kaasschieter	TU Eindhoven	Eindhoven
35.	Bulent	Karasözen	Middle-East TU	Ankara, Turkey
36.	Jan	Kok	CWI	Amsterdam
37.	Hans	Kraaijevanger	Shell SIEP-RTS	Rijswijk
38.	Konstantin	Laevsky	TU Eindhoven	Eindhoven
39.	Debby	Lanser	CWI	Amsterdam
40.	Boris	Lastdrager	CWI	Amsterdam
41.		van der Linden	TU Eindhoven	Eindhoven
42.	Walter	Lioen	CWI	Amsterdam
43.	Jos	Maubach	TU Eindhoven	Eindhoven
44.	Arnold	Mestelaar	Univ. Twente	Enschede
45.	Ellen	Meijerink	Univ Utrecht	Utrecht
46.	Hennie	ter Morsche	TU Eindhoven	Eindhoven
47.	Seva	Nefedov	TU Eindhoven	Eindhoven

48.	Peter	Oswald	Bell Labs	Murray Hill NJ
49.	Johan	Romate	Shell SRTCA	Amsterdam
50.	Dirk	Roose	KU Leuven	Leuven
51.	W.H.A.	Schilders	Philips Research	Eindhoven
52.	Ruud	Schotting	TU Delft	Delft
53.	Jo	Simoens	KU Leuven	Leuven
54.	Ben	Sommeijer	CWI	Amsterdam
55.	Marc	Spijker	RU Leiden	Leiden
56.	Rob	Stevenson	KU Nijmegen	Nijmegen
57.	Andrew	Stuart	Stanford University	Stanford CA
58.	Jacques	de Swart	CWI / Paragon DT	Amsterdam
59.	J.H.M.	ten Thije Boonkkamp	TU Eindhoven	Eindhoven
60.	C.R.	Traas	Univ. Twente	Enschede
61.	Geert	Uytterhoeven	KU Leuven	Leuven
62.	Denis	Vanderstraeten	KU Leuven	Leuven
63.	Stefan	Vandewalle	KU Leuven	Leuven
64.	Wolter	van der Veen	MacNeal–Schwendler	Gouda
65.	Arthur E.P.	Veldman	RU Groningen	Groningen
66.	Menno	Verbeek	Univ Utrecht	Utrecht
67.	Kees	Verhoeven	TU Eindhoven	Eindhoven
68.	Jan	Verwer	CWI	Amsterdam
69.	K.	Wang	TU Eindhoven	Eindhoven
70.	Ivo	Wenneker	TU Delft	Delft
71.	Piet	Wesseling	TU Delft	Delft
72.	Paul M.	de Zeeuw	CWI	Amsterdam
73.	Mario	Zennaro	Università di Trieste	Trieste Italia
74.	Paul	Zegeling	Univ Utrecht	Utrecht
75.	Barbara	Zubik-Kowal	RU Leiden	Leiden