A BRIEF SURVEY ON THE NUMERICAL DYNAMICS FOR FUNCTIONAL DIFFERENTIAL EQUATIONS

Barnabas M. Garay
Department of Mathematics
Budapest University of Technology and Economics
H-1521 Budapest, Hungary

June 14, 2004

Keywords: delay equations, Runge–Kutta discretizations, invariant manifolds.

GYULA FARKAS (1972–2002) IN MEMORIAM

Abstract

This is a survey on discretizing delay equations from a geometric–qualitative view–point. Concepts like compact attractors, hyperbolic periodic orbits, the saddle structure around hyperbolic equilibria, center–unstable manifolds of equilibria, inertial manifolds, structural stability, and Kamke monotonicity are considered. Error estimates for smooth and nonsmooth initial data in various C^j topologies are provided. The emphasis is put on Runge–Kutta methods with natural interpolants. The paper ends with a collection of the related results on retarded functional differential equations with bounded delay.

1 Introduction

The first monograph on numerical methods for delay differential equations has been published by [Bellen & Zennaro, 2003]. Together with their own, they present relevant results of Baker, Brunner, Enright, Guglielmi, Hayashi, Hairer, in't Hout, Iserles, Jackiewicz, Koto, Maset, Tavernini, Torelli, Vermiglio and many other researchers. The book is almost 400 pages long and the bibliography contains 288 items.

The principle of organizing the material in [Bellen & Zennaro, 2003] is that — mutatis mutandis — all discretization methods used for ordinary differential equations can be applied for delay and neutral equations as well. Results for ordinary differential equations are followed by those on delay and neutral equations. Both similarities and differences (compared to the case of ordinary differential equations) are analysed in details. In line with the mainstream tradition of presenting numerical analysis for ordinary diferential equations [Butcher, 1987], [Hairer, Norsett & Wanner, 1993], the emphasis is put on convergence and stability properties of Runge–Kutta methods. The technicalities esp. those related to the choice of the stepsize sequence depend on the type of the delay crucially. Delay equations of the form $\dot{x}(t) = f(t, x(t), x(t-\tau))$ and neutral equations of the form $\dot{x}(t) = f(t, x(t), x(t-\tau), \dot{x}(t-\tau))$ ($t \geq t_0$, $t_0 \in \mathbf{R}$) are considered. With increasing complexity, the delay can be constant ($\tau = \tau_0 > 0$), bounded and time dependent ($\tau = \tau(t) \in [0, \tau_0]$) bounded and space dependent ($\tau = \tau(x(t)) \in [0, \tau_0]$), and proportional ($\tau = qt$ with some $q \in (0,1)$ and $t_0 \geq 0$) — initial data are functions defined on the interval $[t_0 - \tau_0, t_0]$ and $[(1-q)t_0, t_0]$, respectively. Multiple and distributed delays are discussed incidentally.

The monograph [Bellen & Zennaro, 2003] has grown out of *traditional numerical analysis*. Of course the authors are well aware of the fact that the phase space of a delay equation is a function space. The approximating solution is computed first at the mesh points and then, via interpolation, on the intervals between mesh points. However, little attention is paid to the question if geometric–qualitative aspects of the solution dynamics are preserved under discretization. The main object of investigation is the relation between individual solution trajectories and their numerical approximation in $\mathbf{R}^n$. In other words, the much younger *tradition of numerical dynamics*, i.e. of handling numerical methods from the view-point of dynamical system theory plays a rather limited role in [Bellen & Zennaro, 2003].

Thus it is not without any reason to write a brief survey on delay equations placed within the general framework of numerical dynamics. When doing this, we reconsider some central topics discussed in [Stuart & Humphries, 1996], the first comprehensive presentation of numerical dynamics for ordinary differential equations:

- discretization methods as approximating semidynamical systems,
- compact attractors
- hyperbolic periodic orbits
- stable and unstable manifolds of hyperbolic equilibria.

Throughout this paper, we consider only autonomous equations with bounded delay and focus our attention to Runge–Kutta methods with polynomial interpolation. Aspects of

- inertial manifolds,

- structural stability,
- Kamke monotonicity

are also discussed. We refer frequently to papers of our late colleague *Gyula Farkas* who died in a car accident on February 27, 2002 — he was to receive his PhD Diploma at the end of the same week.

The development of numerical dynamics started for ordinary and parabolic partial differential equations simultaneously. Retarded equations followed with some delay and were influenced by the corresponding results on parabolic equations as well as by the general theory of semiflows in Banach spaces. An analysis of the related/underlying work on parabolic equations and general semiflows is beyond the scope of this paper. The most important contributions are cited in the literature we refer to. The monograph [Bellen & Zennaro, 2003] surveys connections to the numerics of Volterra integral equations.

2 Discretization as a Family of Approximating Discrete–time Semidynamical Systems

In this Section we collect some basic definitions and results on discretizing functional differential equations. This requires reformulation and restating within the framework of abstract dynamical systems theory. This level of abstractness — which is not needed for describing results on approximating individual solutions in [Bellen & Zennaro, 2003] — is absolutely essential when treating qualitative–geometric phenomena.

2.1 Runge–Kutta discretization for delay equations

For simplicity, take $t_0 = 0$, $\tau_0 = 1$, and consider first the initial value problem

$$\begin{cases} \dot{x}(t) &= f(x(t), x(t-1)) & \text{for } t \geq 0 \\ x(t) &= \eta(t) & \text{for } t \in [-1, 0] \end{cases} \tag{1}$$

where $f : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^n$ is a bounded C^p function with bounded derivatives and $\eta \in C([-1, 0], \mathbf{R}^n)$, the Banach space of contimuous $\mathbf{R}^n$–valued functions on the interval $[0, 1]$. The maximum norm on $C = C([-1, 0], \mathbf{R}^n)$ is denoted by $|| \cdot ||$. The Euclidean norm on $\mathbf{R}^n$ is denoted by $| \cdot |$. The smoothness and boundedness assumptions on f imply that the initial value problem (1) has a unique solution $x = x_{exact}(\cdot, \eta) : [-1, \infty) \to \mathbf{R}^n$. Moreover, formula $(\Phi(t, \eta))(s) = x(t + s, \eta)$, $s \in [-1, 0]$ defines a semidynamical system $\Phi : \mathbf{R}^+ \times C \to C$. With respect to the second variable, Φ is of class C^p, $p = 1, 2, \ldots$. Note that $x_{exact}(\cdot, \eta)$ is differentiable at $t_0 = 0$ if and only if the left–hand side derivative of the initial function η exists at $t_0 = 0$ and $\dot{\eta}(0) = f(\eta(0), \eta(-1))$. It follows immediately that Φ is not differentiable with respect to its first variable. Since (1) defines a nonautonomous ordinary differential equation of the form $\dot{x} = f(x, \eta(t-1))$ on the interval $[0, 1]$, $x_{exact}(\cdot, \eta)|_{[0,1]} = \Phi(1, \eta)$ is of class C^1 and, by a simple induction argument, $x_{exact}(\cdot, \eta)|_{[j,j+1]} = \Phi(j + 1, \eta)$ is of class C^{j+1}, $j = 0, 1, \ldots, p$. Moreover, $x_{exact}(\cdot, \eta)|_{[j,\infty)}$ is of class C^{j+1}, $j = 0, 1, \ldots, p$ — this is the well–known *smoothing property* of the solution semidynamical system.

Fix $h_0 \in (0,1]$ and let $h \in (0,h_0]$. The *stepsize-h explicit Euler discretization operator with piecewise linear interpolant* is defined as $\varphi_{E,PLI} : (0,h_0] \times C \to C$, $(h,\eta) \to \varphi_{E,PLI}(h,\eta)$,

$$(\varphi_{E,PLI}(h,\eta))(s) = \begin{cases} \eta(h+s) & \text{if } s \in [-1,-h] \\ -\frac{s}{h}\eta(0) + \left(1+\frac{s}{h}\right) X_E(h,\eta) & \text{if } s \in [-h,0] \end{cases} \tag{2}$$

where $X_E(h,\eta) = \eta(0) + hf(\eta(0),\eta(-1))$. The right–hand side of formula (2) defines the stepsize-h implicit Euler discretization operator $\varphi_{I,PLI}$ with piecewise linear interpolant if $X_E(h,\eta)$ is replaced by $X_I(h,\eta)$, the unique solution of equation $X = \eta(0)+hf(X,\eta(h-1))$, for $h \in (0,h_1]$, $h_1 > 0$ sufficiently small.

Similarly, every Runge–Kutta method M (known from the numerics of ordinary differential equations [Butcher, 1987], [Hairer, Norsett & Wanner, 1993]) can be applied for equation (1). The *stepsize-h Runge–Kutta discretization operator with piecewise linear interpolant* is defined as $\varphi_{M,PLI} : (0,h_1^*] \times C \to C$, $(h,\eta) \to \varphi_{M,PLI}(h,\eta)$,

$$(\varphi_{M,PLI}(h,\eta))(s) = \begin{cases} \eta(h+s) & \text{if } \ s \in [-1,-h] \\ -\frac{s}{h}\eta(0) + \left(1+\frac{s}{h}\right) X_M(h,\eta) & \text{if } \ s \in [-h,0] \end{cases} \tag{3}$$

where

$$X_M(h,\eta) = \eta(0) + h\sum_{i=1}^{\nu} b_i\, f\left(X^i, \eta(c_i h - 1)\right) \tag{4}$$

with

$$X^i = \eta(0) + h\sum_{j=1}^{\nu} a_{ij}\, f\left(X^j, \eta(c_j h - 1)\right), \quad i = 1,2,\ldots,\nu\,. \tag{5}$$

Here the positive integer ν and the real constants $\{a_{ij}\}_{i,j=1}^{\nu}$, $\{b_i\}_{i=1}^{\nu}$ and $\{c_i\}_{i=1}^{\nu}$ are the parameters of the Runge–Kutta method M. We leave them unspecified but assume that $c_i \in [0,1]$ for $i = 1,2,\ldots,\nu$. Note that for h sufficiently small, say $h \le h_1^* (\le h_0)$, the right-hand side of (5) defines a contraction operator on $\mathbf{R}^n \times \mathbf{R}^n \times \ldots \times \mathbf{R}^n$ (ν times).

The *stepsize-h Runge–Kutta discretization operator with a standard interpolant* is defined as $\varphi_{M,NFI} : (0,h_1^*] \times C \to C$, $(h,\eta) \to \varphi_{M,NFI}(h,\eta)$,

$$(\varphi_{M,NFI}(h,\eta))(s) = \begin{cases} \eta(h+s) & \text{if } \ s \in [-1,-h] \\ \eta(0) + h\sum_{i=1}^{\nu} \beta_i(-\frac{s}{h})\, f\left(X^i, \eta(c_i h - 1)\right) & \text{if } \ s \in [-h,0] \end{cases} \tag{6}$$

where $\{X^i\}_{i=1}^{\nu}$ is determined by (5) and the polynomials $\beta_i : [0,1] \to \mathbf{R}$ satisfy $\beta_i(0) = b_i$, $\beta_i(1) = 0$, and $\beta_i(1-c_j) = a_{ij}$, $i,j = 1,2,\ldots,\nu$. Note that the collection of the requirements on $\{\beta_i\}_{i=1}^{\nu}$ is equivalent to the collection of the properties

$$(\varphi_{M,NFI}(h,\eta))(-h) = \eta(0) \quad , \quad (\varphi_{M,NFI}(h,\eta))(0) = X_M(h,\eta) \quad \textit{and}$$

$$(\varphi_{M,NFI}(h,\eta))(-h + c_i h) = X^i, \quad i = 1,2,\ldots,\nu\,.$$

The first two letters in NFI refer to the terminus technicus "natural" and "of the first class", respectively. Throughout of this paper, we assume that our Runge–Kutta method when applied for ordinary differential equations (like $\dot{x} = f(x,\eta(t-1))$ on the interval $[0,1]$) is of order p.

In most practical implementations, the approximating solution $x_{approx}(\cdot,\eta) : [-1,\infty) \to \mathbf{R}^n$ is computed only at the mesh points $\{k/N\}_{k\geq 0}$. When keeping track on internal stage values, one arrives at the finite sequence of points $\{x_{approx}((k+c_i)/N,\eta)\}_{k\geq 0; i=1,\ldots,\nu} \subset \mathbf{R}^n$. In particular, $x_{approx}(1/N,\eta) = X_M(h,\eta)$ and $x_{approx}(c_i/N,\eta) = X^i$ from (5)–(6), $i = 1,2,\ldots,\nu$.

It is immediate that $\varphi_{M,NFI} : (0,h_1^*] \times C \to C$ is continuous. With respect to the second variable, $\varphi_{M,NFI}$ is of class C^p. For $h \in (0,h_1^*]$ fixed, the iterates $\{\varphi_{M,NFI}^k(h,\eta))\}_{k=0}^{\infty}$ define a discrete–time semidynamical system on C. Our uniformity assumptions on f imply that, for any time $T \geq 1$ and for any ball $B \subset C$, the set

$$\{\varphi_{M,NFI}^k(h,\eta) \in C \mid 1 \leq kh \leq T\ ,\ k \in \mathbf{N}\ ,\ h \in (0,h_1^*]\ ,\ \eta \in B\ \}$$

consists of uniformly bounded and uniformly Lipschitz continuous functions. Note that the set $\Phi([1,T],B)$ consists of uniformly bounded and uniformly Lipschitz continuous functions, too — this is the well–known *compactifying property* of the solution semidynamical system.

From the view–point of a qualitative theory of discretizations, it is natural to define stepsize–h discretization operators as above i.e., as self–maps of the infinite–dimensional function space C. However, this is not quite satisfactory for practical purposes. In practice the initial function $\eta \in C$ is not always explicitly given but only its values on a uniform mesh are known. This leads to a parallel, more practical framework of establishing an abstract theory for discretizations.

Fix a positive integer N. By letting $\Pi_{1/N}(\eta)$ to be the piecewise linear continuous function with vertices $\{(-1+j/N,\eta(-1+j/N)\}_{j=0}^N$, a linear projection $\Pi_{1/N} : C \to C$ is defined. The range of $\Pi_{1/N}$ is denoted by $C_{1/N} \subset C$. Obviously, $C_{1/N}$ can be identified with $\mathbf{R}^{n(N+1)}$ via the linear isomorphism $\eta \to \{\eta(-1+j/N)\}_{j=0}^N$. Thus the stepsize–$1/N$ explicit Euler discretization method when applied to the delay equation (1) can be understood as a mapping $\varphi(1/N,\cdot) = \varphi_{\mathcal{P},E,PLI}(1/N,\cdot) : C_{1/N} \to C_{1/N}$ defined by

$$(\varphi(1/N,\cdot))\left(\left\{\eta\left(\frac{j-N}{N}\right)\right\}_{j=0}^N\right) = \left(\eta\left(\frac{1-N}{N}\right),\eta\left(\frac{2-N}{N}\right),\ldots,\eta(0),X_E(\eta)\right). \qquad (7)$$

Here the lower index $\mathcal{P}$ is an abbreviation for "practical". Note that the sequence of iterates $\{\varphi_{E,PLI}^k(1/N,\eta))\}_{k=0}^{\infty} \subset C$ depends solely on (f and) $\{\eta(-1+j/N)\}_{j=0}^N$, the restriction of η to the finite collection of the mesh points $\{-1+j/N\}_{j=0}^N$ in $[-1,0]$. By definition,

$$\varphi_{E,PLI}^k(1/N,\eta) = \varphi_{\mathcal{P},E,PLI}^k(1/N,\eta) \qquad \textit{whenever} \quad k \geq N \quad \textit{and} \quad \eta \in C\,.$$

A similar construction is possible for general Runge–Kutta methods and leads to the definition of *stepsize–$1/N$ practical Runge–Kutta discretization operators $\varphi_{\mathcal{P},M,PLI}$ with piecewise linear interpolant.*

2.2 Error estimates for smooth initial functions

For ordinary differential equations on a finite time interval $[0,T]$, it is well–known that stepsize–h Runge–Kutta approximating/discretized solutions converge to the exact solution as $h \to 0$. The order of a Runge–Kutta method M refers to the order of this convergence process

(a) on the set of the mesh points $\{kh\}_{k\geq 0}$ in $[0,T]$

(b) on the set of the stage points $\{kh\}_{k\geq 0} \cup \{(k+c_j)h\}_{k\geq 0;\, j=1,\ldots,\nu}$ in $[0,T]$

and, in case method M is combined with an interpolation operator $INTRP$,

(c) on the entire interval $[0,T]$.

The corresponding orders are called the classical (or nodal), the stage and the uniform order, respectively. A separate order can be defined for the interpolation operator $INTRP$ as well.

The monograph [Bellen & Zennaro, 2003] discusses all the order concepts above in the context of delay equations thoroughly. For delay equations of the form (1), their main result goes back to [Bellen, 1984] and can be restated as follows.

Lemma 1 *Let $\varphi_{M,NFI}$ be a Runge–Kutta discretization operator with standard interpolant. Given any finite interval $[0,T]$ and any C^p initial function η, there exists a positive constant K (depending only on f, M, T, as well as on $||\eta'||, \ldots, ||\eta^{(p)}||$) such that*

$$|(\Phi(k/N,\eta))(0) - (\varphi^k_{M,NFI}(1/N,\eta))(0)| \leq K \cdot h^p \tag{8}$$

whenever $0 \neq N, k \in \mathbf{N}$, $k/N \leq T$ and $1/N \leq h_1^*$. Under some additional assumptions on the interpolant NFI, also inequalities

$$||\frac{d^j}{ds^j}\,\Phi(k/N,\eta) - \frac{d^j}{ds^j}\,\varphi^k_{M,NFI}(1/N,\eta)|| \leq K_j \cdot h^{q+1-j} \quad ,\ j=0,1,\ldots,q \tag{9}$$

hold true. Here $q \leq p-1$ stays for the order of the interpolant NFI and constant K_j depends also on the underlying interpolation operator, $j=0,1,\ldots,q$. Derivatives at the mesh points in (9) are meant in the left/right sense.

Suppose that $\eta^{(p)}$ does not exists at some $s=-1+\Delta$ with $h_0<\Delta<1$ but η is C^p on the interval $[-1,-1+\Delta]$. Then the local approximation error satisfies inequality

$$|x_{exact}(h,\eta) - X_M(h,\eta)| \leq const(f,M,\eta)\cdot h^{p+1} \qquad for\ \ h\in(0,\Delta] \tag{10}$$

where $const(f,M,\eta)$ depends only on f, M, as well as on the bounds for $|\eta'|,\ldots,|\eta^{(p)}|$. The restriction "for $h\in(0,\Delta]$" in (10) (which means that the inequality in (10) is not necessarily satisfied on the interval $(0,h_0]$) has important consequences to mesh point selection. In order to have local $\mathcal{O}(h^{p+1})$ error estimates, it implies that Δ has to be chosen for a mesh point. In view of the smoothing property of Φ, a similar argument shows that the mesh (still in order to have local $\mathcal{O}(h^{p+1})$ error estimates) should contain the points $0,1,\ldots,p$ and also the points $1+\Delta, 2+\Delta$ etc. In general this leads to nonuniform mesh with a variable stepsize sequence $(h_1,h_2,\ldots)$, $0<h_m\leq h_0$, $m=1,2,\ldots$. The discretization operator $\varphi_{M,NFI}$ gives rise to one with variable stepsize sequence by defining $\varphi_{M,NFI}(0,\eta)=\eta$ and then, inductively

$$\varphi_{M,NFI}(h_m,\ldots,h_1;\eta) = \varphi_{M,NFI}(h_m,\varphi_{M,NFI}(h_{m-1},\ldots,h_1;\eta)) \quad for \quad m=1,2,\ldots\ .$$

Chapters 4 and 6 of [Bellen & Zennaro, 2003] contain several generalizations of what we called Lemma 1 above even for equations with state–dependent delay as well as for certain types of neutral equations where (still in order to have local $\mathcal{O}(h^{p+1})$ error estimates)

stepsize selection is subject to various constraints. Also these results can be restated within the framework of a nonautonomous dynamical system theory. The "additional assumptions on the interpolant NFI" from Lemma 1 (which go back to [Zennaro, 1986] and constitute one of the mostly involved part of [Bellen & Zennaro, 2003]) are discussed in Subsection 5.2.2. It is a challenging task to find such an interpolation operator that preserves the order of convergence in Lemma 1 and makes $x_{approx}(\cdot,\eta)$ to be of class C^j on the interval $[j,\infty)$, $j=0,1,\ldots,p$. The monograph [Bellen & Zennaro, 2003] refers to several results into this direction but none of them seems to ensure the same smoothness improvement for $x_{approx}(\cdot,\eta)$ along the intervals $\{[j,j+1]\}_{j=0}^{p}$ which is shared by $x_{exact}(\cdot,\eta)$.

There are various pro and contra arguments for variable stepsizes sequences. A major pro argument has already been discussed. Though conflicting with higher order local approximation error estimates, further pro arguments are those behind adaptive error control in [Bellen & Zennaro, 2003], Chapter 7. The major contra argument for variable stepsize sequences is the obvious pro argument for the uniform mesh $\{k/N\}_{k\geq 0}$ we outlined in the two last paragraphs of Subsection 2.1.

However, despite of all efforts of putting stepsizes selection and the error control mechanism on a firm mathematical basis, heuristical aspects can hardly be avoided. This is particularly exemplified by considering a delay equation of the form

$$\dot{x}(t)=f(x(t),x(t-\rho),x(t-1)) \qquad where \quad h_0<\rho<1\,.$$

Our first candidate is the uniform mesh $M_U=\{k/N\}_{k\geq 0}$. In order to go on with the explicit Euler method at a mesh point $k_0/N\leq T$, two earlier values of the approximate solution x_{approx} (at $k_0/N-\rho$ and $k_0/N-1$ given or computed previously) are needed. It follows that the values of η at each point of the set

$$H_T=\{k/N-\ell\rho\in[-\rho,0]\mid k\in\mathbf{Z}\,,\ \ell\in\mathbf{N}\,,\ k/N\leq T\}$$

are also needed. This is a pledoyee for interpolating and working within the $\varphi_{\mathcal{P},E,PLI}$ framework but, especially on moderate time intervals $[0,T]$, also the choice of M_U+H_T (the algebraic sum of the two discrete sets M_U and H_T) as for a new, nonuniform mesh $M_{NU}=M_U+H_T$ is reasonable.

The subsection concludes with an example showing that local error estimates between exact and approximate solutions cannot be uniform in C. Nevertheless, it indicates that, on certain natural subsets of C, uniform error estimates can be expected.

Example 1 *If function f does not depend on the first n coordinates, then (1) simplifies to*

$$\left\{\begin{array}{lcll}\dot{x}(t) & = & f(x(t-1)) & \text{for } t\geq 0\\ x(t) & = & \eta(t) & \text{for } t\in[-1,0]\,.\end{array}\right. \tag{11}$$

Suppose that $f(0)=0$ and $\eta(-jh)=0$ whenever $h=1/N$ and $j=0,1,\ldots,N$. Then $(\Phi(1,\eta))(s)=\int_{-1}^{s}f(\eta(u))\,du$ but $(\varphi^N_{E,PLI}(1/N,\eta))(s)=0$ for each $s\in[-1,0]$. In particular, $||\Phi(1,\eta)-\varphi^N_{E,PLI}(1/N,\eta)||$ can be arbitrarily large. Note that in our case,

$$||\Phi(1,\eta)-\varphi^N_{E,PLI}(1/N,\eta)||\leq\int_{-1}^{0}|f(\eta(u))|\,du\leq\mathcal{L}\int_{-1}^{0}|\eta(u)|\,du$$

where $\mathcal{L}$ stays for the Lipschitz constant of f. Note that $||\Phi(1,\eta) - \varphi^N_{E,PLI}(1/N,\eta)|| \to 0$ as $N \to \infty$ for each $\eta \in C$ because on $[0,1]$ ((1) is equivalent to a nonautonomous ordinary differential equation and thus) (11) simplifies to the integration problem $x(t) - \eta(0) = \int_0^t f(\eta(u-1))\,du$).

2.3 Error estimates for nonsmooth initial functions

If the initial function $\eta \in C$ is Lipschitz with constant $Lip(\eta) \le L$, then the local approximation error satisfies inequality

$$|x_{exact}(h,\eta) - X_M(h,\eta)| \le const(f,M,L)\cdot h^2 \qquad for\ \ h \in (0,h_1^*]\,. \tag{12}$$

This is a consequence of (10) when applied to a sequence of C^1 Lipschitz functions $\{\eta_k\}_{k=1}^\infty \subset C$ with the properties that $Lip(\eta_k) \to Lip(\eta)$ and $||\eta_k - \eta|| \to 0$ as $k \to \infty$.

Lemma 2 *Let $\varphi_{M,NFI}$ be a Runge–Kutta discretization operator with standard interpolant. Given any finite time interval $[0,T]$ and a finite collection of Lipschitz initial functions $\eta, \xi_1, \dots \xi_j$ with constants $Lip(\eta), Lip(\xi_1), \dots, Lip(\xi_j) \le L$ and $||\xi_1||, \dots, ||\xi_j|| \le 1$, the derivatives of the approximation error (as a j–linear operator between $C \times C \times \dots \times C$ (j times) and C) satisfy the $j = 0, 1, \dots, p-1$ chain of inequalities*

$$||[\frac{d^j}{d\eta^j}\,\Phi(k/N,\eta)](\xi_1,\dots,\xi_j) - [\frac{d^j}{d\eta^j}\,\varphi^k_{M,NFI}(1/N,\eta)](\xi_1,\dots,\xi_j)|| \le \kappa_j/N \tag{13}$$

whenever $0 \ne N, k \in \mathbf{N}$, $k/N \le T$ and $1/N \le h_1^*$. The positive constant κ_j, $j = 0, 1, \dots, p-1$ depends only on f, M, T, L, and on the underlying interpolation operator.

Proof. Case $j = 0$ is a direct consequence of inequality (12) via the standard Gronwall argument [Hairer, Norsett & Wanner, 1993].

In order to prove case $j = 1$, we pass to a somewhat higher level of abstractness. Still with the initial value problem (1) in mind, we use the standard notation from the theory of retarded functional differential equations [Hale, 1977] and write $\dot x(t) = g(x_t)$, $x_0 = \eta$ instead. We consider also the initial value problem $\dot y(t) = [g'(x_t)]y_t$, $y_0 = \xi$ for the first variational equation. Define

$$G\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} g(x_t) \\ [g'(x_t)]y_t \end{pmatrix} \qquad \text{and} \qquad \Theta\left(h, \begin{pmatrix} \eta \\ \xi \end{pmatrix}\right) = \begin{pmatrix} \Phi(t,\eta) \\ [\frac{d}{d\eta}\,\Phi(t,\eta)]\xi \end{pmatrix}$$

where $t \ge 0$ and $\Phi : \mathbf{R}^+ \times C \to C$ denotes the solution semidynamical system for equation $\dot x(t) = g(x_t)$. It is clear that $\Theta : \mathbf{R}^+ \times (C \times C) \to C \times C$ is the solution semidynamical system for the retarded functional differential equation

$$\begin{pmatrix} \dot x(t) \\ \dot y(t) \end{pmatrix} = G\begin{pmatrix} x_t \\ y_t \end{pmatrix} \quad , \quad t \ge 0\,. \tag{14}$$

Similarly, with $\varphi : (0,h_1^*] \times C \to C$ denoting an "approximation operator" for Φ, define

$$\psi\left(h, \begin{pmatrix} \eta \\ \xi \end{pmatrix}\right) = \begin{pmatrix} \varphi(h,x) \\ [\frac{d}{d\eta}\,\varphi(h,\eta)]\xi \end{pmatrix} \quad \text{whenever} \ \ h \in (0,h_1^*] \ \text{ and } \ \begin{pmatrix} \eta \\ \xi \end{pmatrix} \in C \times C\,.$$

Suppose that $\varphi = \varphi_{M,NFI,g}$ comes from a Runge–Kutta method with interpolant. Then the very same Runge–Kutta method applies to equation (14) and gives rise to operator $\varphi_{M,NFI,G}$. Analysing (5)–(6), it is not hard to show that $\psi = \varphi_{M,NFI,G}$. We arrived at the conclusion that case $j = 1$ of (13) follows from inequality (12) — when applied to equation (14) instead of $\dot{x}(t) = g(x_t)$ — via the standard Gronwall argument. (Unfortunately, the uniformity assumptions we imposed on g remain no longer valid for G. The second coordinate of G is unbounded when $||y_t|| \to \infty$. However, if our interest is reduced to a bounded subset of $C \times C$, then no difficulties arise.)

The remaining cases $j = 2, \ldots, p-1$ follow by induction. *Q.E.D.*

Since $\Phi(0, \eta) = id_C \eta = \eta$ for each $\eta \in C$ and $\frac{d^j}{d\eta^j}\, \Phi(\cdot, \eta)$ is differentiable, we obtain immediately from (13) — or, by a direct analysis of the definition — that

$$||[\frac{d^j}{d\eta^j}\, \varphi_{M,NFI}(1/N, \eta)](\xi_1, \ldots, \xi_j) - [\frac{d^j}{d\eta^j}\, id_C](\xi_1, \ldots, \xi_j)|| \leq \tilde{\kappa}_j / N \tag{15}$$

whenever $1/N \leq h_1^*$ and the initial functions $\eta, \xi_1, \ldots \xi_j$ are Lipschitz with constants $Lip(\eta)$, $Lip(\xi_1)$, ..., $Lip(\xi_j) \leq L$ and $||\xi_1||, \ldots, ||\xi_j|| \leq 1$, $j = 0, 1, \ldots, p-1$. Of course the positive constant $\tilde{\kappa}_j$ depends only on f, M, L, and on the underlying interpolation operator, $j = 0, 1, \ldots, p-1$.

Recall that Lipschitz functions in C are of bounded variation and that the total variation $Totv(\eta)$ is not greater than $Lip(\eta)$.

Hence our next result is an improvement over Lemma 2 for the explicit/implicit Euler method with piecewise linear interpolant.

Lemma 3 *Consider only the special cases $\varphi_{M,NFI} = \varphi_{E,PLI}$ or $\varphi_{I,PLI}$. Given any finite time interval $[0, T]$ and a finite collection of initial functions $\eta, \xi_1, \ldots \xi_j$ of bounded variation with $Totv(\eta)$, $Totv(\xi_1)$, ..., $Totv(\xi_j) \leq L$, the $j = 0, 1, \ldots, p-1$ chain of inequalities (13) still holds true.*

Proof. Applying the standard Gronwall argument [Hairer, Norsett & Wanner, 1993] we have already referred to, we point out first that

$$|(\Phi(1, \eta))(0) - (\varphi_{E,PLI}^N(1/N, \eta))(0)| \leq 2^{-1} e^{\mathcal{L}_1} (B + 4n\mathcal{L}_2 \cdot Totv(\eta))/N \tag{16}$$

where B be is an upper bound for $|f|$ on $\mathbf{R}^n \times \mathbf{R}^n$ and $\mathcal{L}_i$ stays for the Lipschitz constant of f with respect to the i–th variable, $i = 1, 2$. On the time interval $[0, 1]$, (1) is equivalent to the initial value problem $\dot{x} = f(x, q(t))$, $x(0) = x_0 = \eta(0)$ where $q(t) = \eta(t-1)$. Let $\Psi(\cdot; t_*, x_*) : [t_*, 1] \to \mathbf{R}^n$ denote the right–hand side solution to the nonautonomous ordinary differential equation $\dot{x} = f(x, q(t))$ with initial data $(t_*, x_*) \in [0, 1] \times \mathbf{R}^n$. For brevity, we write $h = 1/N$,

$$t_k = kh \quad , \quad x_k = (\varphi_{E,PLI}^N(1/N, \eta))(t_k - 1) \qquad for \quad k = 0, 1, \ldots, N\,,$$

and $a = \mathcal{L}_1 B h^2/2$, $b = \mathcal{L}_2$, $d = e^{\mathcal{L}_1 h}$. Finally, for $k = 0, 1, \ldots, N-1$, define

$$E_k = |\Psi(t_k; 0, x_0) - x_k| \qquad and \qquad c_k = \int_{t_k}^{t_{k+1}} |q(u) - q(t_k)|\, du$$

and observe that the right–hand side of inequality (16) is equal to $E_N = |\Psi(t_N; 0, x_0) - x_N|$.

We claim that $E_{k+1} \leq E_k d + a + bc_k$ for each $k = 0, 1, \ldots, N-1$. In fact, we have for $k = 0, 1, \ldots, N-1$ by the triangle inequality that

$$E_{k+1} \leq |\Psi(t_{k+1}; t_k, \Psi(t_k; 0, x_0)) - \Psi(t_{k+1}; t_k, x_k)| + |\Psi(t_{k+1}; t_k, x_k) - (x_k + hf(x_k, q(t_k))| \, .$$

The first term can be estimated by using Gronwall lemma. In fact, for each $t \in [t_k, t_{k+1}]$, we have that

$$|\Psi(t; t_k, \Psi(t_k; 0, x_0)) - \Psi(t; t_k, x_k)|$$

$$= |\Psi(t_k; 0, x_0) + \int_{t_k}^{t} f(\Psi(u; t_k, \Psi(t_k; 0, x_0)), q(u))\, du - (x_k + \int_{t_k}^{t} f(\Psi(u; t_k, x_k), q(u))\, du|$$

$$\leq |\Psi(t_k; 0, x_0) - x_k| + \mathcal{L}_1 \int_{t_k}^{t} |\Psi(u; t_k, \Psi(t_k; 0, x_0)) - \Psi(u; t_k, x_k)|\, du)$$

and, a fortiori, the first term is not greater then $e^{\mathcal{L}_1 h} E_k$. On the other hand, the second term is bounded by

$$|x_k + \int_{t_k}^{t} f(\Psi(u; t_k, x_k), q(u))\, du - (x_k + \int_{t_k}^{t_{k+1}} f(x_k, q(t_k))\, du)|$$

$$\leq \mathcal{L}_1 \int_{t_k}^{t} |\Psi(u; t_k, x_k) - x_k|\, du + \mathcal{L}_2 \int_{t_k}^{t_{k+1}} |q(u) - q(t_k)|\, du$$

and the claim follows from observing that $x_k = \Psi(t_k; t_k, x_k)$ and $\Psi(\cdot; t_k, x_k)$ is Lipschitz with constant $\leq B$.

Starting from $E_0 = 0$, we conclude easily by induction that

$$E_N \leq a\, \frac{d^N - 1}{d - 1} + b \sum_{k=0}^{N-1} c_k d^{N-k-1} \leq 2^{-1} e^{\mathcal{L}_1} (Bh + \mathcal{L}_2 \sum_{k=0}^{N-1} c_k) \, .$$

It remains to prove that $\sum_{k=0}^{N-1} c_k \leq 2nh \cdot Totv(\eta)$. In fact, by the Jordan decomposition theorem, every coordinate function of q can be represented as $q_i = v_i - w_i$ where v_i and w_i are monotone increasing continuous real functions on $[0, 1]$ with the property that

$$Totv(v_i)\, ,\, Totv(w_i) \leq Totv(q_i) \leq Totv(q) \qquad i = 1, \ldots, n \, .$$

It follows immediately that

$$\sum_{k=0}^{N-1} c_k \leq \sum_{k=0}^{N-1} \sum_{i=1}^{n} \int_{t_k}^{t_{k+1}} (|v_i(u) - v_i(t_k)| + |w_i(u) - w_i(t_k)|)\, du$$

$$\leq \sum_{i=1}^{n} \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} (|v_i(t_{k+1}) - v_i(t_k)| + |w_i(t_{k+1}) - w_i(t_k)|)\, du$$

$$\leq \sum_{i=1}^{n} h \cdot (Totv(v_i) + Totv(w_i)) \leq 2nh \cdot Totv(q) = 2nh Totv(\eta) \, .$$

(As a direct consequence of the uniform continuity of η, note that $\sum c_k \to 0$ as $N \to \infty$.)

By a repeated use of the standard Gronwall argument (when combined with piecewise linear interpolation), case $j = 0$ of the $\varphi_{M,NFI} = \varphi_{E,PLI}$, $L = Totv(\eta)$ version of inequality (13) follows with $\kappa_0 = const \cdot e^{\mathcal{L}_1 T}(1 + L)$ easily.

The $\varphi_{M,NFI} = \varphi_{I,PLI}$ case can be reduced to the $\varphi_{M,NFI} = \varphi_{E,PLI}$ case already proven. The crucial point is to show that

$$|(\varphi^N_{I,PLI}(1/N, \eta))(0) - (\varphi^N_{E,PLI}(1/N, \eta))(0)| \leq const \cdot (1 + Totv(\eta))/N \, .$$

Without referring to the Jordan decomposition theorem any more, this follows via a simplified version of the recursion we used in deriving (16) above. (Having only Theorem 4.B in mind, we did not check if (16) holds true for a general discretization operator $\varphi_{M,NFI}$.)

The proof of the remaining cases $j = 1, 2, \ldots, p-1$ is the same as in the proof of Lemma 2. *Q.E.D.*

Both for ordinary and delay differential equations, one–step and multistep methods, several versions of inequalities (8), (9), (13), and (15) are known from the literature. A weaker version of Lemma 3 has been stated in [Garay & Lóczi, 2004]. Inequality (16) is new.

The definition of an abstract discretization operator for equation (1) as well for the more abstract equation $\dot{x}(t) = g(x_t)$ we refer to in Section 4 below are based on case $j = 0, 1$ of inequalities (13) and (15).

In contrast to inequalities (8) and (9) which concern an individual exact and an individual approximating trajectory, the $j \geq 1$ cases of inequalities (13) and (15) relate to a collection of exact and approximating trajectories. Qualitative theory cannot live without differentiating with respect to initial data. This is why C^1 inequalities like case $j = 1$ of (13) (estimating the difference between exact and approximating solutions in C^1 topologies on the phase space) play a fundamental role in almost all papers on numerical dynamics. The typical result is that, for stepsizes sufficiently small, hyperbolic orbit configurations are preserved by discretization.

For ordinary differential equations, an abstract definition for discretization operators is based on C^j properties. In one of the earliest papers on the qualitative theory of discretizations, [Beyn & Lorenz, 1987] suggest the following definition. Consider an ordinary differential equation $\dot{x} = f(x)$ where $f : \Omega \to \mathbf{R}^n$ is a C^{p+r+1} function. Let S be a compact subset of Ω. A mapping $\varphi : (0, h_0] \times S \to \mathbf{R}^n$ is an *abstract discretization operator of order* p if

(i) φ admits a C^{p+r+1} extension to an open neighborhood of $[0, h_0] \times S$ in $\mathbf{R} \times \Omega$

(ii) $|\Phi(h, x) - \varphi(h, x)| \leq K h^{p+1}$ whenever $(h, x) \in (0, h_0] \times S$

(iii) φ is locally determined by f. In other words, there exists a continuous function $\Delta : [0, h_0] \to \mathbf{R}^+$ with the properties that $\Delta(0) = 0$ and, for all $(h, x) \in (0, h_0] \times S$, $\varphi(h, x)$ is determined by the restriction of f to the set $\{z \in \mathbf{R}^n \mid |z - x| \leq \Delta(h)\}$.

As a simple consequence of assumptions (i)–(ii), $\varphi(h, \cdot)$ is a C^{p+r+1} diffeomorphism of S onto $\varphi(h, S)$ for h sufficiently small, and — on condition that $\Phi((k-1)h, x) \in S$ and $\varphi^{k-1}(h, x) \in S$ —

$$|\frac{d^j}{dx^j} \, \Phi(kh, x) - \frac{d^j}{dx^j} \, \varphi^k(h, x)| \leq \kappa_j(T) \cdot h^{min\{p, p+r-j\}} \quad , \; j = 0, 1, \ldots, p + r \tag{17}$$

whenever $k \in \mathbf{N}$, $h \in (0, h_0]$, $kh \leq T$ and $x \in S$. Clearly Runge–Kutta methods are subject to assumptions (i)–(iii). Moreover, for Runge–Kutta methods, inequality (17) is satisfied if f is chosen from the less smoother class of C^{p+r} functions. Mutatis mutandis, assumptions (i)–(iii) make sense if f is defined on a compact smooth manifold $\mathcal{M}$. This leads to the definition of abstract discretization operators on $\mathcal{M}$. For stepsize h small enough, $\varphi(h, \cdot)$ is a C^{p+r+1} self–diffeomorphism of $\mathcal{M}$. Also the C^j inequality (17) remains valid in the manifold setting. For details, see [Li, 1997], [Garay, 2001].

Thus $\{\varphi(h, \cdot)\}_{h \in (0,h_0]}$ is a one–parameter family of diffeomorphisms approximating the one–parameter family of time–h diffeomorphism $\{\Phi(h, \cdot)\}_{h \in (0,h_0]}$ of the continuous–time solution dynamical system $\Phi : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n$ (or locally, $\Phi : \mathbf{R} \times \Omega \hookrightarrow \mathbf{R}^n$; or $\Phi : \mathbf{R} \times \mathcal{M} \to \mathcal{M}^n$). Consequently, for $h \in (0, h_0]$ fixed, discretization theory is part of perturbation theory for discrete–time dynamical systems. However, with $h \to 0$, both $\varphi(h, \cdot)$ and $\Phi(h, \cdot)$ approach the identity, an operator which behaves badly in perturbation theory: The behaviour of h as of a small parameter is not entirely regular. We conclude that *the proof of a qualitative result in discretization theory requires a thorough reconsideration of the proof of the underlying abstract perturbation result in discrete dynamics* with stepsize h as an additional small parameter, *and the derivation of the accompanying error estimates.*

3 Qualitative Numerics for Delay Equations

What we described in the last paragraph for ordinary differential equations remains valid for delay equations, too. However, one is confronted with two major difficulties. These are the lack of uniform local error estimates and the lack of backward solvability. Fortunately, for L large enough say $L \geq L_*$, the closed set

$$C_{Lip(L)} = \{\eta \in C \mid \eta \text{ is Lipschitz with constant } Lip(\eta) \leq L\}$$

is positively invariant with respect to the exact as well as to the discretized dynamics. By (13) and (15), C^j estimates on $C_{Lip(L)}$ are uniform. Badly enough, $C_{Lip(L_*)}$ is nowhere dense in C. What really helps is the smoothing/compactifying property of the exact and the discretized dynamics. As for the asymptotic theory, it implies that dynamical systems in finite and semidynamical systems in infinite dimension can be treated in a parallel way [Hale, 1988]. Distinquished subsets of the phases space like unstable manifolds of hyperbolic equilibria or of hyperbolic periodic orbits, inertial manifolds, and compact attractors consist of full trajectories i.e., of trajectories defined on the entire real line $\mathbf{R}$. In particular, compact attractors and certain kinds of invariant manifolds of delay equations belong to $C_{Lip(L_*)}$. This is why, in a final analysis, their qualitative discretization properties are (almost) the same as of their counterparts in ordinary differential equations.

From now on, let $p \geq 2$ and assume that all the regularity conditions we imposed on (1) in Subsection 2.1 are satisfied.

3.1 The simplest hyperbolic orbit configurations

The three major objects of the phase space investigated in [Stuart & Humphries, 1996] on numerical ordinary differential equations are

- compact attractors (i.e. asymptotically stable compact invariant sets)
- hyperbolic periodic orbits
- hyperbolic equilibria, together with their stable and unstable manifolds

In a well–defined technical sense, compact attractors, the saddle structure about hyperbolic equilibria, and periodic orbits are only slightly perturbed under discretization. As for compact attractors, the hyperbolic structure preserved is the transversal intersection structure between trajectories near the attractor and the level surfaces of suitable Liapunov functions. (The dynamics within the attractor itself is not assumed to be hyperbolic and can be changed dramatically under discretizaton.) The presentation in [Stuart & Humphries, 1996] is based on the original papers [Kloeden & Lorenz, 1986], [Beyn, 1987a], [Beyn, 1987b]. No doubt these three papers belong to those few marking the birth of numerical dynamics as an independent field of research in the late eighties.

In what follows we present the corresponding results for delay equations.

Theorem 1 *Consider the delay equation* $\dot{x}(t) = f(x(t), x(t-1))$ *and let* $\varphi_{M,NFI} : (0, h_1^*] \times C \to C$ *be a Runge–Kutta discretization operator with standard interpolant.*

- *A.) [Kloeden & Schropp, 2004]: Let* $\emptyset \neq \mathcal{A}$ *be a compact attractor for the continuous-time solution semidynamical system* Φ*. Then, for stepsize* $1/N$ *sufficiently small, the discrete-time semidynamical system* $\varphi_{M,NFI}(1/N, \cdot)$ *has a nonempty compact attractor* $\mathcal{A}_{1/N}$ *and the limiting process* $\mathcal{A}_{1/N} \to \mathcal{A}$ *as* $N \to \infty$ *(both in* $C_{Lip(L_*)}$ *with nice Liapunov estimates and consequently, by using the general attraction results in Chapter 2 of [Hale, 1988], also in* C*) is upper semicontinuous.*

- *B.1.) [In't Hout & Lubich, 1998]: Let* Γ *be an exponentially stable periodic orbit for the continuous-time solution semidynamical system* Φ*. Then, for stepsize* $1/N$ *sufficiently small, the discrete-time semidynamical system* $\varphi_{M,NFI}(1/N, \cdot)$ *has an exponentially stable invariant curve* $\Gamma_{1/N}$ *(both in* $\mathcal{N} \cap C_{Lip(L_*)}$ *and in* $\mathcal{N}$ *where* $\mathcal{N}$ *is a suitable neighborhood of* Γ *in* C *but in the second case constant* κ *in the estimate* $d(\varphi^k_{M,NFI}(1/N, \eta), \Gamma_{1/N}) \leq \kappa \cdot \mu^{k/N}$ *depends on* η*;* $\mu < 1$ *is fixed) such that* $d_{Hausdorff}(\Gamma, \Gamma_{1/N}) \leq const/N^p$*.*

- *B.2.) [Farkas, 2003]: Let* Γ *be a hyperbolic periodic orbit for the continuous–time solution semidynamical system* Φ *and assume that the period of* Γ *is at least two (i.e. two times the delay). Then, for stepsize* h *sufficiently small (and not only for* $h = 1/N$ *with* N *large), the discrete–time semidynamical system* $\varphi_{M,NFI}(h, \cdot)$ *has a hyperbolic invariant curve* Γ_h *such that in normal coordinates around* Γ*, both* $|\Gamma_h|$ *and* $Lip(\Gamma_h)$ *are of order* h*.*

In line with (17) and the general estimates for discretized normally hyperbolic compact invariant manifolds of ordinary differential equations [Garay, 2001], it seems plausible in Parts B1) and B2) that $\Gamma_{1/N} = \mathcal{F}_{1/N}(\Gamma)$ where $\mathcal{F}_{1/N}$ is a C^{p+r} embedding of Γ into $\mathbf{R}^n$ and the norm distance in $C^j(\Gamma, \mathbf{R}^n)$ between $\mathcal{F}_{1/N}$ and the inclusion of Γ in $\mathbf{R}^n$ is of order $1/N^{min\{p, p+r-j\}}$, $j = 0, 1, \ldots, p+r$.

The unpublished PhD dissertation *Gyula Farkas: On numerical dynamics of functional differential equations, Budapest University of Technology, 2002*, contains a lower semicontinuity result for discretized compact attractors of delay equations, the analogue of the one in [Stuart & Humphries, 1996] from the theory of discretized ordinary differential equations. As for upper semicontinuity, Farkas refers to [Gedeon & Hines, 1999] on upper semicontinuity of Morse sets under explicit ODE–Euler discretization of a one–dimensional delay equation (which results in a cyclic feedback system of ordinary differential equations). Though conceptually much easier, we note that Theorem 1.A is not a consequence of the results in [Gedeon & Hines, 1999].

For the rest of this subsection, we assume that $f(0,0) = 0$ or, equivalently, that $\eta_0 = 0 \in C$ is an equilibrium for Φ. The next result starts with a center–unstable versus strongly–stable $C = CU \times SS$ product decomposition of the phase space (invariant with respect to the linear semidynamical system generated by the solutions of the linearized equation $\dot{y}(t) = f'_x(0,0)y(t) + f'_y(0,0)y(t-1)$ and determined by its characteristic equation). Thus the equilibrium $0 \in C$ is not necessarily hyperbolic. As a consequence of basic spectral decomposition theory, note that the linear subspace CU is of finite dimension.

Theorem 2 *[Farkas, 2002a]: Consider the delay equation $\dot{x}(t) = f(x(t), x(t-1))$ again and let $\varphi_{\mathcal{P},E,PLI}(1/N, \cdot) : C_{1/N} \to C_{1/N}$ be the stepsize–$1/N$ practical explicit Euler discretization operator with piecewise linear interpolant. Assume that the equilibrium point $0 \in C$ has a center–unstable manifold of the form Graph(G), where $G : CU \to SS$ is a C^2 function. Then for N large enough (and under very mild additional technical conditions) $C_{1/N}$ admits a center–unstable versus strongly–stable $C_{1/N} = CU_{1/N} \times SS_{1/N}$ product decomposition with the properties as follows. Operator $\varphi_{\mathcal{P},M,PLI}(1/N, \cdot)$ has an invariant manifold of the form Graph$(G_{1/N})$, where $G_{1/N} : CU_{1/N} \to SS_{1/N}$ is a C^2 function. In addition, there exists a linear isomorphism $P_{1/N} : CU \to CU_{1/N}$ such that, for $j = 0$ and $j = 1$,*

$$\| \frac{d^j}{d\eta^j} \Pi_{1/N} G \; - \; \frac{d^j}{d\eta^j} G_{1/N} P_{1/N} \| \to 0 \qquad as \quad N \to \infty \, . \tag{18}$$

Together with Graph(G), also Graph$(G_{1/N})$ is exponentially attractive, with asymptotic phase depending continuously on the stepsize as $N \to \infty$.

Theorem 2 in [Farkas, 2002a] is accompanied by C^2 existence and C^1 approximation results for exact and discretized stable manifolds that correspond to $\bar{S}$ in the center–unstable versus stable product structure $CU \times \bar{S}$ where $\bar{S}$ is the finite–dimensional invariant subspace determined by a bounded set of the roots of the characteristic equation lying to the left of those belonging to CU. In the case of hyperbolic equilibria, also a numerical Grobman–Hartman lemma for partial linearizations [Farkas, [2001b] as well as C^1 shadowing results [Farkas, 2002a] are given.

As a preparation for the next subsection, we recall the simplest ordinary differential equation result on numerical structural stability [Garay, 1996]. The numerical saddle structure results in [Beyn, 1987a] can be interpreted as follows: Given a hyperbolic equilibrium $x_0 \in \mathbf{R}^n$ of an ordinary differential equation, there exist a neighborhood $\mathcal{U}$ of x_0 in $\mathbf{R}^n$, a constant $\kappa > 0$ and, for each $h \in (0, h_1^*]$, there exists a homeomorphism $\mathcal{H}_h$ of $\mathcal{U}$ into $\mathbf{R}^n$ with the properties that

$$\mathcal{H}_h(x_0) = x_0 \qquad and \qquad |\mathcal{H}_h(x) - x| \leq \; \kappa h^p \quad \text{for each} \quad x \in \mathcal{U} \tag{19}$$

and, last but not least,

$$\mathcal{H}_h(\Phi(h,x)) = \varphi(h,\mathcal{H}_h(x)) \qquad whenever \quad x \in \mathcal{U}\ ,\ \Phi(h,x) \in \mathcal{U}\ . \tag{20}$$

In other words, in the vicinity of hyperbolic equilibria, the exact and the discretized dynamics are conjugate and discretization is nothing else but an almost–identical coordinate transformation. With $\mathcal{H}_h$ being a C^{p+r} diffeomorphism, note that (19) and (20) can be proved in the vicinity of nonequilibria, too [Garay & Simon, 2001].

3.2 Inertial manifolds and structural stability

Throughout this subsection, we restrict ourselves to a certain type of delay equations with small delay. The smallness of the delay seems to be necessary for the C^2 smoothness of the inertial manifold. (The existence of C^1 inertial manifolds can be proved with moderate delay. However, if the delay is not small, then the gap condition (which is the basis for proving higher order smoothness) is violated and, for the time being, there is no way out of this difficulty. For details and references, see [Robinson, 1999], [Farkas, 2002b], [Farkas, 2002c], [Chicone, 2003]. Here we restrict ourselves to reminding the reader that inertial manifolds are global center–unstable invariant manifolds.) On the other hand, the C^2 smallness of the inertial manifold is necessary to apply [Li, 1997] on numerical structural stability for ordinary differential equations in proving Part B.) of the Theorem below. All numerical structural stability results we are aware of require at least C^2 smoothness assumptions.

Recall the definition of the practical stepsize–$1/N$ explicit Euler discretization operator $\varphi_{\mathcal{P},E,PLI}(1/N,\cdot) : C_{1/N} \to C_{1/N}$ from the last paragraph of Subsection 2.1. If the delay is $\varepsilon > 0$, then the phase space is $C^\varepsilon = C([-\varepsilon,0],\mathbf{R}^n)$. A trivial modification of formula (7) gives rise to the definition of the stepsize–ε/N practical explicit Euler discretization operator $\varphi^\varepsilon_{\mathcal{P},E,PLI}(1/N,\cdot) : C^\varepsilon_{1/N} \to C^\varepsilon_{1/N}$.

Theorem 3 *Consider the delay equation $\dot{x}(t) = Ax(t) + a(x(t)) + b(x(t-\varepsilon))$ where A is an $n \times n$ real matrix, $a, b : \mathbf{R}^n \to \mathbf{R}^n$ are bounded C^2 functions with bounded derivatives, and ε is a positive parameter.*

- *A.) [Farkas, 2002c]: Then there exists an $\varepsilon_0 > 0$ with the properties as follows. For every $\varepsilon \in (0,\varepsilon_0]$, the delay equation has an invariant manifold of the form $Graph(J^\varepsilon)$ where $J^\varepsilon : CU^\varepsilon \to SS^\varepsilon$ is of class C^2, the linear subspace CU^ε is finite–dimensional, and the $C^\varepsilon = CU^\varepsilon \times SS^\varepsilon$ product decomposition is given by $CU^\varepsilon = \pi^\varepsilon(C^\varepsilon)$, $SS^\varepsilon = (id|_{C^\varepsilon} - \pi^\varepsilon)(C^\varepsilon)$ with $\pi^\varepsilon : C^\varepsilon \to C^\varepsilon$, $(\pi^\varepsilon(\eta))(s) = e^{As}\eta(0)$, $s \in [-\varepsilon,0]$. In addition, for N large enough (and under very mild additional technical conditions), $C^\varepsilon_{1/N}$ admits a $C^\varepsilon_{1/N} = CU^\varepsilon_{1/N} \times SS^\varepsilon_{1/N}$ product decomposition with the properties as follows. Operator $\varphi^\varepsilon_{\mathcal{P},E,PLI}(1/N,\cdot)$ has an invariant manifold of the form $Graph(J^\varepsilon_{1/N})$, where $J^\varepsilon_{1/N} : CU^\varepsilon_{1/N} \to SS^\varepsilon_{1/N}$ is a C^2 function. In addition, there exists a linear isomorphism $P^\varepsilon_{1/N} : CU^\varepsilon \to CU^\varepsilon_{1/N}$ such that for $j = 0$ and $j = 1$,*

$$||\frac{d^j}{d\eta^j}\,\Pi^\varepsilon_{1/N} J^\varepsilon \;-\; \frac{d^j}{d\eta^j}\,J^\varepsilon_{1/N} P^\varepsilon_{1/N}|| \to 0 \qquad as \quad N \to \infty\ . \tag{21}$$

Together with $Graph(J^\varepsilon)$, also $Graph(J^\varepsilon_{1/N})$ is exponentially attractive, with asymptotic phase depending continuously on the stepsize as $N \to \infty$.

- *B.) [Farkas, 2002c]:* (CONTINUATION.) *Assume, in addition, that the solution flow* $\Psi : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n$ *of the limiting ordinary differential equation* $\dot{x} = Ax + a(x) + b(x)$ *is structurally stable and that the point at the* $\{\infty\}$ *of* $\mathbf{R}^n$ *is repulsive. Then, for* N *large enough, there exists a homeomorphism* $\mathcal{H}^{\varepsilon}_{1/N}$ *of* $\mathbf{R}^n$ *onto* $Graph(J^{\varepsilon}_{1/N})$ *and a continuous time–reparametrization mapping* $r^{\varepsilon}_{1/N} : \mathbf{R}^n \to \mathbf{R}^+$ *such that*

$$\mathcal{H}^{\varepsilon}_{1/N}(\Psi(r^{\varepsilon}_{1/N}(x), x)) = \varphi^{\varepsilon}_{\mathcal{P},E,PLI}(1/N, H^{\varepsilon}_{1/N}(x)) \qquad whenever \quad x \in \mathbf{R}^n \,. \tag{22}$$

If Ψ *is Morse–Smale and gradient–like, then* $r^{\varepsilon}_{1/N}(x) = 1/N$ *for each* $x \in \mathbf{R}^n$.

The reader is asked to make a comparison between (18) and (21) as well as between (20) and (22).

The proof of Theorem 3 requires a very careful handling of standard inertial manifold techniques like spectral decomposition, manipulations with cut–off functions on finite–dimensional subspaces, fixed–point equations in weighted sequences of Banach spaces, fiber contraction theorem etc. extended for discretizations. As for numerical structural stability, it is just an application of the fundamental theorem on numerical structural stability in [Li, 1997], derived as by–product of the Moser–Robbin–Robinson approach to Smale's structural stability theorem.

3.3 Kamke monotonicity

Assume that, for some constant $\gamma \geq 0$, condition

$$\begin{cases} (f_i)'_{x_j}(x,y) \;\geq\; \gamma & \text{if } (x,y) \in \mathbf{R}^n \times \mathbf{R}^n,\ i,j = 1,2,\ldots,n \text{ and } i \neq j, \\ (f_i)'_{y_j}(x,y) \;\geq\; \gamma & \text{if } (x,y) \in \mathbf{R}^n \times \mathbf{R}^n,\ i,j = 1,2,\ldots,n \end{cases} \tag{23}$$

Here of course f_i stands for the i–th coordinate function of f, further $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ denote the first n and the last n coordinate variables of f_i, $i = 1, 2, \ldots, n$, respectively. By letting $x \leq \widetilde{x}$ for $x, \widetilde{x} \in \mathbf{R}^n$ if and only if $x_i \leq \widetilde{x}_i$ for each $i = 1, 2, \ldots, n$, a closed partial order on $\mathbf{R}^n$ is defined. The closed partial order $\leq$ on $\mathbf{R}^n$ generates a closed partial order $\preceq$ on C. In particular, $\eta \preceq \widetilde{\eta}$ holds if and only if $\eta(s) \leq \widetilde{\eta}(s)$ for each $s \in [-1, 0]$. As an easy consequence of assumption (20) the semi–dynamical system Φ is Kamke monotone [Smith, 1995]. In other words, inequality $\Phi(t, \eta) \preceq \Phi(t, \widetilde{\eta})$ holds true whenever $t \geq 0$ and $\eta, \widetilde{\eta} \in C$ with $\eta \preceq \widetilde{\eta}$.

Theorem 4 *Consider the initial value problem (1) under condition (23). Then*

- *A.) [Garay & Lóczi, 2004]: Let* $\gamma > 0$. *Given any Runge–Kutta method* M *satisfying* $b_i \geq 0$ *for* $i = 1, \ldots, \nu$, *the discretization operator* $\varphi_{M,PLI}$ *(with piecewise linear interpolation) is monotone in the sense that, for sufficiently small stepsize* h *and for any initial functions with* $\eta \preceq \widetilde{\eta}$, *also the order relation* $\varphi_{M,PLI}(h, \eta) \preceq \varphi_{M,PLI}(h, \widetilde{\eta})$ *holds true.*

- *B.) [Garay & Lóczi, 2004]: Let* $\gamma > 0$. *Suppose that* $f_i(x, y) \geq 0$ *and* $(f_i)'_{x_i}(x, y) \geq 0$ *for* $i = 1, 2, \ldots, n$. *Then, for sufficiently small stepsize* h *and for any nondecreasing initial function* η, *we have that* $\varphi_{E,PLI}(h, \eta) \preceq \Phi(h, \eta) \preceq \varphi_{I,PLI}(h, \eta)$.

- *C.) [Kloeden & Schropp, 2003]: Let $\gamma = 0$. Suppose we are given a Runge–Kutta method M with the properties that $a_{ij} > 0$ for $i, j = 1, \ldots, \nu$ and $b_i > 0$ for $i = 1, \ldots, \nu$. Then the discretization operator $\varphi_{M,PLI}$ is monotone. Moreover, the positivity assumption on the Runge–Kutta matrix $A = \{a_{ij}\}_{i,j=1}^{\nu}$ can be weakened to the nonnegativity assumption on the matrix function $\tau \to (I + \tau A)^{-1} A$, required on some nondegenerate τ–interval $[0, \tau_0]$.*

An iterative combination of Parts A.) and B.) formulates and generalizes the well–known observation that, given a one–dimensional ordinary differential equation with all solutions convex, then every solution curve is above the broken line determined by the explicite, and under the broken line determined by the implicite Euler method. Part C.) is entirely of different character. In a strong resemblance to results on contractivity in numerical ordinary differential equations [Hairer, Norsett & Wanner, 1993], it provides a suffient condition for a Runge–Kutta method to preserve monotonicity of the solution dynamics under discretization. In the light of the elegant counterexamples in [Kloeden & Schropp, 2003], this sufficient condition is almost necessary.

Even in the numerical contexts of differential equation theory, the word "monotonicity" can be used in a number of various ways. Monotonicity of iterative methods for delay equations has already been investigated in [Erbe & Liu, 1991].

4 Remarks on Functional Differential Equations

The previous considerations suggest that all the Theorems above are valid for retarded functional differential equations of the form $\dot{x}(t) = g(x_t)$ where $x_t(s) = x(t+s)$ for $s \in [0, 1]$, and $g : C \to \mathbf{R}^n$ is of class C^p, $p \geq 2$. Gyula Farkas has always formulated his results in this more general framework. However, occasionally, he carried out the proofs only for the special case $g(x_t) = f(x(t), x(t-1))$ and indicated the technical modifications needed for a general g. He considered assumptions (2)-(4) in [Farkas, 2003] and assumptions (i)-(vi) of his Lemma 8 in [Farkas, 2002c] as a general definition of discretization operators for equations of the form $\dot{x}(t) = g(x_t)$ and $\dot{x}(t) = Lx_t + g(x_t)$, respectively. Here $L : C \to \mathbf{R}^n$ is a bounded linear operator (which, by a theorem of Riesz, can be represented as a Stieltjes integral $L\eta = \int_{-1}^{0} \eta(s)\, d\vartheta(s)$).

The general definitions above have little relevance to practical purposes. For example, in all numerical implementations we are aware of, $L\eta$ is replaced by a finite sum like $\sum_{i=1}^{N} N^{-1} \left(\vartheta((1-j)/N) - \vartheta(-j/N)\right) \eta(-j/N)$. Several references on the numerics of retarded functional differential equations — chosen in the spirit of the Bellen–Zennaro monograph — are contained in [Maset, 2003].

The numerics of equations with infinite delay — partly because of the depth of the underlying functional analysis — is more complicated. There are only sporadic results into this direction [Liu, 1997]. We cite also the papers [Koto, 1999], [Insperger & Stépán, 2002] representing those devoted to some qualitative aspects of numerical bifurcation and numerical stability of retarded/delay equations. We are not aware of any papers on the numerics of delay equations with computer–assisted proofs.

All in all, we conclude by emphasizing that the large gap that characterized the relation of abstract dynamical systems theory and the numerical practice of solving differential equa-

tions until the early nineties of the last century, has been considerably filled in the last ten years. Having read papers like [Shub, 1986] or [Matijasevich, 1985] on numerical methods, it is clear to us that many of the most distinguished mathematicians have (i) hoped for (ii) guessed (iii) foreseen (iv) worked for this development. Among them John-von-Neumann is pioneer number one.

Acknowledgements

Parts of this paper were written during a stay of the author at the University of Padova. Hospitality of the Department of Mathematics is greatfully acknowledged. The author is indebted to Wolf-Jürgen Beyn, Giovanni Colombo, Peter Kloeden, and Johannes Schropp for valuable discussions during the preparation of the paper. The paper is supported by the Hungarian National Science Foundation OTKA No. T037491.

References

Bellen, A. [1984] "One-step collocation for delay differential equations," *J. Comput. Appl. Math.* **10**, 275–283.

Bellen, A. & Zennaro, M. [2003] *Numerical Methods for Delay Differential Equations* (Oxford University Press, Oxford).

Beyn, W. J. [1987a] "On the numerical approximation of phase portraits near stationary points," *SIAM J. Num. Anal.* **24**, 1095–1113.

Beyn, W. J. [1987b] "On invariant closed curves of one-step methods," *Numer. Math.* **51**, 103–122.

Beyn, W. J. & Lorenz, J. [1987] "Center manifolds of dynamical systems under discretization," *Num. Funct. Anal. Optimiz.* **9**, 318–414.

Butcher, J. F. [1987] *The Numerical Analysis of Ordinary Differential Equations* (Wiley, London).

Chicone, C. [2003] "Inertial and slow manifolds for delay equations with small delays," *J. Differ. Eq.* **190**, 364–406.

Erbe, L. & Liu, X. [1991] "Monotone iterative methods for differential systems with finite delay," *Appl. Math. Comput.* **43**, 43–64.

Farkas, G. [2001a] "Unstable manifolds for RFDEs under discretization: the Euler method," *Computers Math. Applic.* **42**, 1069–1081.

Farkas, G. [2001b] "A Grobman–Hartman result for retarded functional differential equations with an application to the numerics of hyperbolic equilibria," *Z. angew. Math. Phys.* **52**, 421–432.

Farkas, G. [2002a] "A numerical C^1-shadowing result for retarded functional differential equations," *J. Comput. Appl. Math.* **45**, 269–289.

Farkas, G. [2002b] "Nonexistence of uniform exponential dichotomies for delay equations," *J. Differ. Eq.* **182**, 266–268.

Farkas, G. [2002c] "Small delay inertial manifolds under numerics: a numerical structural stability result," *J. Dyn. Differ. Eq.* **14**, 549–588.

Farkas, G. [2003] "Discretizing hyperbolic periodic orbits of delay differential equations," *Z. angew. Math. Mech.* **83**, 38–49.

Garay, B. M. [1996] "On structural stability of ordinary differential equations with respect to discretization methods," *Numer. Math.* **72**, 449–479.

Garay, B. M. [2001] "Estimates in discretizing normally hyperbolic compact invariant manifolds of ordinary differential equations," *Comput. Math. Applic.* **42**, 1103–1122.

Garay, B. M. & Simon, P.L. [2001] "Numerical flow-box theorems under structural assumptions," *IMA J. Numer Anal.* **21**, 733–749.

Garay, B. M. & Lóczi, L. [2004] "Monotone delay equations and Runge–Kutta discretizations," *Funct. Differ. Eq.* **11**, 59–67.

Gedeon, T. & Hines, G. [1999] "Upper semicontinuity of Morse sets of a discretization of a delay-differential equation," *J. Differ. Eq.* **151**, 36–78.

Hairer, E., Norsett, S. P. & Wanner, G. [1993] *Solving Ordinary Differential Equations I. Nonstiff Problems* (Springer, Berlin).

Hale, J. K. [1977] *Theory of Functional Differential Equations* (Springer, Berlin).

Hale, J. K. [1988] *Asymptotic behaviour of dissipative systems* (AMS, Providence).

Insperger, T. & Stépán, G. [2002] "Stability chart for the delayed Mathieu equation," *Royal Soc. London Proc. Ser. A. Math. Phys. Eng. Sci.* **458**, 1989–1998.

In't Hout, K. & Lubich, Ch. [1998] "Periodic orbits of delay differential equations under discretization," *BIT* **38**, 72–91.

Kloeden, P. E. & Lorenz, J. [1986] "Stable attracting sets in dynamical systems and their one-step discretization," *SIAM J. Num. Anal.* **23**, 986–995.

Kloeden, P. E. & Schropp, J. [2003] "Runge–Kutta methods for monotone differential and delay equations," *BIT* **43**, 571–586.

Kloeden, P. E. & Schropp, J. [2004] "Stable attracting sets in delay differential equations and in their Runge–Kutta discretization," (submitted)

Koto, T. [1999] "Neumark–Sacker bifurcation in the Euler method for a delay differential equations," *BIT* **39**, 110–115.

Li, M.C. [1997] "Structural stability of flows under numerics," *J. Differ. Eq.* **141**, 1–12.

Liu, Y. [1997] "On the θ-method for delay equations with infinite time lag," *J. Comput. Appl. Math.* **71**, 177–190.

Maset, S. [2003] "Numerical solution of retarded functional differential equations as abstract Cauchy problems," *J. Comput. Appl. Math.* **16**, 259–282.

Matijasevich, Yu. V. [1985] "A posteriori interval analysis," *EUROCAL* Vol. 2., Caviness, B.F. (Ed.) (Springer, Berlin), pp. 328–334.

Robinson, J. C. [1999] "Inertial manifolds with and without delay," *Discete Cont. Dyn. Systems* **5**, 813–824.

Shub, M. [1986] "Some remarks on dynamical systems and numerical analysis," *Dynamical systems and Partial Differential Equations* Lara–Carrero L. & Lewowicz, J. (Eds.) (Univ. Simon Bolivar, Caracas), pp. 69–91.

Smith, H.L. [1995] *Monotone Dynamical Systems* (AMS, Providence).

Stuart, A.M. & Humphries, A.R. [1996] *Dynamical Systems and Numerical Analysis* (Cambridge University Press, Cambridge).

Zennaro, M. [1986] "Natural continuous extensions of Runge–Kutta methods," *Math. Comp.* **46**, 119–133.