DYNAMICS OF NUMERICS OF LINEARIZED COLLOCATION METHODS

MELUSI KHUMALO

Dynamics of Numerics of Linearized Collocation Methods

by

Melusi Khumalo

A thesis submitted to the School of Graduate Studies in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Mathematics and Statistics
Memorial University of Newfoundland

December, 1997

St John's Newfoundland

ABSTRACT

Many ordinary differential equations that describe physical phenomena possess solutions that cannot be obtained in closed form. To obtain the solutions to these systems, the use of numerical schemes is unavoidable. Traditional numerical analysis concerns itself with obtaining error bounds within finite closed time intervals; however, the study of asymptotic or long term behaviour of solutions generated by numerical schemes has attracted a lot of interest in recent years. It is now well established that numerical schemes for nonlinear autonomous differential equations can admit asymptotic solutions which do not correspond to those of the ODE.

This thesis studies linearized one-point collocation methods, contributing to this important investigation by considering bifurcation phenomena in autonomous ODEs and studying the dynamics of the methods for nonautonomous ODEs.

Using the theory of normal forms, it is established that the common codimension-1 bifurcations that exist in continuous dynamical systems will occur in the methods at the same phase space location. However, the methods can exhibit period doubling bifurcations, which are necessarily spurious. They also introduce a singular set, which drastically affects the global dynamics of the methods.

The technique of stroboscopic sampling of the numerical solution is used to study the dynamics of nonautonomous ODEs with periodic solutions, and conditions under which the methods have a unique periodic solution that is asymptotically stable, are stated explicitly. A link between these conditions and nonautonomous linear and nonlinear stability theory is established.

ACKNOWLEDGEMENTS

I would first like to express my apppreciation to my supervisor, Dr. E.A.D. Foster, for his unfailing support and encouragement. I would also like to thank the other members of my supervisory committee, Dr. H. Brunner and Dr. R. Charron for their valuable input, the Department for giving me the opportunity and providing pleasant facilities and a friendly environment in which to work, and the Canadian Commonwealth Scholarship and Fellowship Program (administered by the International Council for Canadian Studies) for financial assistance.

Not much could be accomplished without the moral support and encouragement I have always received from my family, church and many friends. I wish to express a heartfelt thank you to all.

Contents

1 Preliminaries **1**
- 1.1 Introduction . . . 1
- 1.2 Autonomous Case . . . 2
- 1.3 Invariant Sets and Manifolds . . . 12
- 1.4 Nonautonomous Case . . . 17
- 1.5 Linearized Collocation Methods . . . 19
  - 1.5.1 Derivation of One-Point Collocation Methods . . . 19
  - 1.5.2 Linearization . . . 21
- 1.6 Outline of Thesis . . . 23

2 History and Background **25**
- 2.1 Introduction . . . 25
- 2.2 Spurious Fixed Points . . . 27
- 2.3 Characterization of Regularity in Runge-Kutta Methods . . . 28
- 2.4 Study of Spurious Fixed Points . . . 29
- 2.5 Bifurcations to Spurious Period-2 Solutions . . . 31
- 2.6 Branches of Spurious Solutions . . . 33
- 2.7 Global Asymptotic Behaviour . . . 34
- 2.8 Relevance of Thesis . . . 36

3 Autonomous Equations **38**
- 3.1 Introduction . . . 38
- 3.2 Codimension-1 Bifurcations . . . 39
  - 3.2.1 Saddle-Node Bifurcation . . . 40
  - 3.2.2 Transcritical Bifurcation . . . 48
  - 3.2.3 Pitchfork Bifurcation . . . 52
  - 3.2.4 Period Doubling Bifurcation . . . 56
  - 3.2.5 Conclusion . . . 62

3.3 Hopf and Neimark-Sacker Bifurcations 63
3.3.1 Bifurcation Values in Differential Equations 63
3.3.2 Bifurcation Values in the Numerical Methods 64
3.4 Examples 68
3.4.1 Van der Pol's Oscillator (Unforced) 68
3.4.2 Forced Oscillators and the Method of Averaging 69

4 Spurious Poles in Numerical Solutions to Autonomous Equations 75
4.1 Introduction 75
4.2 Normal Form Equations: Codimension-1 Bifurcations 76
4.2.1 Saddle-node Bifurcation 76
4.2.2 Transcritical Bifurcation 78
4.2.3 Subcritical Pitchfork Bifurcation 80
4.2.4 Remarks 82
4.3 Motivation: A Predator-Prey System 83
4.3.1 Linearized One-point Collocation Methods 84
4.3.2 Basins of Attraction 87
4.4 Spurious Pole-type Behaviour 91
4.4.1 Remarks 94

5 Nonautonomous Equations and Periodic Solutions 96
5.1 Introduction 96
5.1.1 Dynamical Systems Approach 98
5.2 Linear Case 99
5.2.1 Linear Case with $a(t) = 0$ 100
5.2.2 Linear Case with $a(t) = -1$ 104
5.2.3 Linear Case with $a(t) = -1 + \epsilon\rho(t)$ 111
5.2.4 Conditional AN-stability and AN-stability 118
5.3 Nonlinear Case 121
5.3.1 Linearized One-Point Collocation Methods 123
5.3.2 The Dynamical Systems Approach 124
5.3.3 Nonlinear Stability Theory 130
5.3.4 Discussion 134
5.4 Conclusion 135

6 **Higher Dimensional Linear Systems** **137**
6.1 Introduction 137
6.2 Linearized Collocation Methods 139
6.2.1 Periodic Solutions 140
6.2.2 Remarks 141

7 **Concluding Remarks** **143**
7.1 Summary of Thesis 143
7.1.1 Scalar Parameter-dependent ODEs 143
7.1.2 Planar Parameter-dependent ODEs 144
7.1.3 Planar Non-parameter-dependent ODEs 145
7.1.4 Periodic Solutions - Nonautonomous ODEs 146
7.2 Suggestions for Future Work 147

List of Figures

1.1 Bifurcation diagram: Explicit Euler method, applied to (1.8). . 11

3.1 Saddle-node: $f = \alpha + x^2$, $c_1 > 0.5$ 44
3.2 Saddle-node: $f = \alpha + x^2$, $c_1 < 0.5$ 45
3.3 Transcritical: $f = \alpha x + x^2$, $c_1 > 0.5$ 51
3.4 Transcritical: $f = \alpha x + x^2$, $c_1 < 0.5$ 51
3.5 Supercritical Pitchfork: $f = \alpha x - x^3$, $c_1 < 0.5$ 55
3.6 Supercritical Pitchfork: $f = \alpha x - x^3$, $c_1 > 0.5$ 55
3.7 Period doubling: $f = (\alpha_0 - \alpha)x + x^3$, $\alpha_0 = \frac{2}{h(2c_1-1)}$, $c_1 > 0.5$. 60
3.8 Period doubling: $f = (\alpha_0 - \alpha)x + x^3$, $\alpha_0 = \frac{2}{h(2c_1-1)}$, $0 < c_1 < \frac{1}{6}$ 60
3.9 Period doubling: $f = (\alpha_0 - \alpha)x + x^3$, $\alpha_0 = \frac{2}{h(2c_1-1)}$, $\frac{1}{6} < c_1 < \frac{1}{2}$ 61
3.10 Period doubling: $f = (\alpha_0 - \alpha)x - x^3$, $\alpha_0 = \frac{2}{h(2c_1-1)}$, $\frac{1}{6} < c_1 < \frac{1}{2}$ 61
3.11 Bifurcation diagram of the averaged equations (3.32) 71
3.12 Bifurcation values of collocation methods for the averaged equations of the forced VDP oscillator. 74

4.1 Saddle-node: $f = \alpha + x^2$, $c_1 = 1$, $h = 0.1$ 78
4.2 Transcritical: $f = \alpha x + x^2$, $c_1 = 1$ and $h = 0.1$. 80
4.3 Subcritical Pitchfork: $f = \alpha x + x^3$, $c_1 = 1$, $h = 0.1$ 81
4.4 Bifurcation curve for the fixed point (2.1,1.98). 85
4.5 Bifurcation curve for the fixed point (0,0). 85
4.6 Bifurcation curve for the fixed point (1,0). 86
4.7 Bifurcation curve for the fixed point (3,0). 86
4.8 Basin of attraction – Predator Prey Model. 89
4.9 Numerical basin of local fixed points ($h = 0.1, c_1 = .5$) 90
4.10 Singular set of Linearized One-point collocation methods with $hc_1 = 0.05$. Singular set (- · - ·), single orbit (—). 91

5.1 Numerical results for (5.9) with $c_1 = 1/2$. $k = 4$ ($\times\ \times$), $k = 5$ (—), $k = 10$ ($- - -$) 104
5.2 Numerical results of (5.18) using linearized implicit midpoint method: $k = 3$ (—) and $k = 10$ ($- - -$) 110
5.3 Numerical results of (5.18) using Explicit Euler method: $k = 3$ (—) and $k = 10$ ($- - -$) 111
5.4 Least k for unique periodic solution. Explicit Euler ($- - -$), Linearized Implicit Midpoint (—), Linearized Implicit Euler ($\times\ \times$) 115

Chapter 1

Preliminaries

1.1 Introduction

Let

$$x' = f(t, x), \qquad x(0) = x_0, \tag{1.1}$$

where $f : \mathbb{R} \times I \subset \mathbb{R}^m \to \mathbb{R}^m$, be an ordinary differential equation.

If f does not depend explicitly on t, then the differential equation

$$x' = f(x), \qquad x(0) = x_0, \tag{1.2}$$

where $f : I \subset \mathbb{R}^m \to \mathbb{R}^m$, is said to be *autonomous*, otherwise it is *nonautonomous*.

This thesis will be primarily concerned with various aspects of dynamical behaviour that may exist in the solution of (1.1) and the conditions under which these are exhibited by linearized one-point collocation methods. In this chapter, we will summarize the well known aspects of dynamical behaviour that will be of interest in this work, and derive this class of numerical methods which we will be using in the study.

1.2 Autonomous Case

Assume $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$, so that (1.2) has unique solution, $\Phi(x_0, t)$, defined for $t = (\beta_1, \beta_2)$ with $0 \in (\beta_1, \beta_2)$.

If $f(x^*) = 0$, x^* is said to be an *equilibrium* point of (1.2).

Define the orbits of x_0:

$$\begin{aligned}
\gamma^+(x_0) &= \bigcup_{t\in[0,\beta_2)} \Phi(x_0, t) \\
\gamma^-(x_0) &= \bigcup_{t\in(\beta_1,0]} \Phi(x_0, t) \\
\gamma(x_0) &= \bigcup_{t\in(\beta_1,\beta_2)} \Phi(x_0, t).
\end{aligned}$$

If $\gamma^-(x_0)$ is bounded, we define

$$\alpha(x_0) = \lim_{t\to\beta_1^+} \Phi(x_0, t),$$

and if $\gamma^+(x_0)$ is bounded, define

$$\omega(x_0) = \lim_{t\to\beta_2^-} \Phi(x_0, t).$$

We call $\alpha(x_0)$ the α-limit set of x_0 and $\omega(x_0)$ the ω-limit set of x_0.

We state some stability concepts for equilibrium points in autonomous equations (see Wiggins [32]).

Definition 1.2.1 *An equilibrium point, x^*, of (1.2) is said to be* **Lapunov stable** *if, for any $\epsilon > 0$, there is a $\delta = \delta(\epsilon) > 0$, such that, for every x_0 for which $\|x_0 - x^*\| < \delta$, the solution $\Phi(t, x_0)$ of (1.2) satisfies $\|\Phi(t, x_0) - x^*\| < \epsilon$ for all $t \geq 0$.*

x^* is said to be asymptotically stable if it is stable and, in addition, there is an $r > 0$ such that $\|\Phi(t, x_0) - x^*\| \to 0$ as $t \to +\infty$ for all x_0 satisfying $\|x_0 - x^*\| < r$.

An equilibrium point x^* of $x' = f(x)$ is said to be *hyperbolic* if none of the eigenvalues of $f_x(x^*)$ have zero real part.

Stability of a hyperbolic equilibrium point can be determined from the linearization of the vector field:

Theorem 1.2.2 *Let f be a C^1 function. If all the eigenvalues of the Jacobian matrix, $f_x(x^*)$, have negative real parts, then the equilibrium point x^* is asymptotically stable.*

A hyperbolic equilibrium point is called a *saddle* if some, but not all, the eigenvalues of $f_x(x^*)$ have real parts greater than zero and the rest have real parts less than zero. If all the eigenvalues have negative (positive) real part, then the hyperbolic equilibrium point is called a stable (an unstable) node or a sink (a source). An equilibrium point is called a *spiral* if all the eigenvalues have nonzero real and imaginary parts.

For scalar autonomous equations, the α- and ω-limit sets are equilibrium points, if they exist. If a numerical method is used to discretize a differential equation, then the discrete system so obtained is a dynamical system in its own right — a discrete dynamical system. Since discrete systems possess, in general, much richer dynamics than their continuous counterparts (see Devaney [6] for a comprehensive discussion of the dynamics of discrete systems), we are confronted with a number of possibilities:

- Can the numerical method exhibit limit sets that are not present in *any* continuous system?
- Can the numerical method exhibit limit sets that do not exist in the specific system it has been designed to solve?

In other words, how do the aspects of dynamical behaviour of the numerical method compare with that of the continuous system? Let

$$x \mapsto x + h\phi(x, h) := g(x) \tag{1.3}$$

be the map corresponding to the discretization of (1.2) by a chosen numerical method. Starting with the initial value x_0, the map generates a solution sequence (positive orbit) denoted by

$$\gamma^+(x_0) = \{x_0, g(x_0), g^2(x_0), \dots, g^n(x_0), \dots\}.$$

Then, we have the following definition.

Definition 1.2.3 *A point X^* such that $g(X^*) = X^*$, that is, $\phi(X^*, h) = 0$ in (1.3) is called a* fixed point *of the map.*

Definition 1.2.4 *A fixed point, X^*, of (1.3) is said to be* **stable** *if, for any $\epsilon > 0$, there is a $\delta = \delta(\epsilon) > 0$, such that, for every x_0 for which $\|x_0 - X^*\| < \delta$, the solution sequence originating at x_0 satisfies $\|g^n(x_0) - X^*\| < \epsilon$ for all $n \geq 0$.*

X^* is said to be asymptotically stable if it is stable and, in addition, there is an $r > 0$ such that $\|g^n(x_0) - X^*\| \to 0$ as $n \to +\infty$ for all x_0 satisfying $\|x_0 - X^*\| < r$.

It should be noted that the stability of X^* is dependent on the discretization parameter h.

A fixed point X^* of the map $x \mapsto g(x)$ is hyperbolic if none of the eigenvalues of $g_x(X^*)$ has modulus one.

Stability of a hyperbolic fixed point can also be determined from linearization of the map:

Theorem 1.2.5 *Let g be a C^1 function. If all the eigenvalues of the Jacobian matrix, $g_x(X^*)$, have moduli less than one, then the fixed point X^* is asymptotically stable. If at least one of the eigenvalues has modulus greater than one, X^* is unstable.*

A hyperbolic fixed point is called a *saddle* if some, but not all, the eigenvalues of $g_x(X^*)$ have moduli greater than one and the rest have moduli less than one. If all the eigenvalues have moduli less than one (greater than one), then the hyperbolic fixed point is called a stable (an unstable) node or a sink (source). A fixed point is called a *spiral* if all the eigenvalues have nonzero real and imaginary parts.

The phenomenon of period 2 solutions in discrete dynamics, which we now define, has no counterpart in continuous systems of dimension less than 3 (see Humphries [18] and Stuart & Humphries [31]).

Definition 1.2.6 *A solution sequence of the form $X_{2n} = u^*$, $X_{2n+1} = v^*$, where $u^* \neq v^*$, is called a* period 2 *solution of (1.3).*

If a numerical method generates period 2 solutions, we know that such solutions are necessarily spurious since no continuous systems admit such solutions.

The study of dynamics of numerics has focused primarily on Runge-Kutta and multi-step methods. We present here these two classes of numerical methods.

Consider the general consistent linear k-step method

$$\sum_{j=0}^{k} \alpha_j X_{n+j} = h \sum_{j=0}^{k} \beta_j f(X_{n+j}) \tag{1.4}$$

with fixed $h > 0$. Define the first and second characteristic polynomials by $\rho(z) = \sum_{j=0}^{k} \alpha_j z^j$ and $\sigma(z) = \sum_{j=0}^{k} \beta_j z^j$ respectively. Since the method is consistent, $\rho(1) = 0$ and $\sigma(1) = \rho'(1) = a$. If the method is *zero-stable*, a is a nonzero constant.

We also consider the s-stage Runge-Kutta method

$$X_{n+1} = X_n + h \sum_{i=1}^{s} b_i f(z_i), \tag{1.5}$$

where

$$z_j = X_n + h \sum_{j=1}^{s} a_{ij} f(z_i), \qquad i = 1, 2, \ldots, s. \tag{1.6}$$

Here, we define $A := [a_{ij}]$ and $\mathbf{b} := [b_i]^T$.

Definition 1.2.7 *A function $f : \mathbb{R}^m \to \mathbb{R}^m$ is said to be Lipschitz on $X \subset$*

$\mathbb{R}^m$ *if there exists a number* $L > 0$ *such that*

$$\|f(x) - f(y)\| \leq L\|x - y\| \qquad \text{for all } x, y \in X, \tag{1.7}$$

where $\|\cdot\|$ *is an* $\mathbb{R}^m$ *norm.* L *is called the Lipschitz constant.* f *is globally Lipschitz if* f *is Lipschitz on* $\mathbb{R}^m$*, and locally Lipshitz if* f *is Lipschitz on all bounded subsets of* $\mathbb{R}^m$.

Iserles [19] related the notion of a fixed point of (1.3) with that of an equilibrium point of (1.2) for Runge-Kutta and multi-step methods. He proved that Runge-Kutta schemes can possess *extra* fixed points that are not equilibrium solutions of the differential equation; these, whenever they exist, are said to be *spurious* solutions. Linear multi-step methods do not possess spurious fixed points.

Definition 1.2.8 (Stuart & Humphries [31]) *A numerical method for (1.2) which does not admit spurious fixed points is said to be* **regular of degree 1,** $R^{[1]}$. *A method which is not* $R^{[1]}$ *is* **irregular of degree 1**, *denoted* $IR^{[1]}$.

Definition 1.2.9 (Stuart & Humphries [31]) *A numerical method for (1.2) which does not have period two solutions is said to be* **regular of degree 2,** $R^{[2]}$. *A method which is not* $R^{[2]}$ *is* **irregular of degree 2**, *denoted* $IR^{[2]}$. *A method that is both* $R^{[1]}$ *and* $R^{[2]}$ *is said to be* regular, *denoted* $R^{[1,2]}$.

The explicit Euler method is the only explicit R-K method that is $R^{[1]}$, no explicit R-K method that is $R^{[2]}$ exists, and the highest order of regular Runge-Kutta methods is 2 (Hairer *et al.* [13], Stein [28]). If we would like to use only $R^{[1,2]}$ R-K methods, then we are essentially restricted to low order methods. However, Stein [28] proved that a limit stepsize, h_l, exists, below which a Runge-Kutta method exhibits no spurious fixed points. Also, Humphries [18] proved that if f is globally Lipschitz, spurious fixed points in R-K methods cannot exist for arbitrarily small h. However, Stein demonstrated that the limit stepsize h_l can be computed.

Theorem 1.2.10 (Humphries [18]) *The linear multi-step method (1.4) is not $R^{[2]}$ if $\rho(-1) = 0$. If $\rho(-1) \neq 0$, and the method is zero-stable, then it is $R^{[2]}$ if and only if $\sigma(-1) = 0$.*

The backward differentiation formula (BDF) is an example of an $R^{[2]}$ linear multi-step method.

The following famous example is used to demonstrate that numerical methods can exhibit very complicated dynamical behaviour even if the underlying system is very simple.

Example (Griffiths *et al.* [10])

Consider the well-known logistic equation

$$x' = \alpha x(1-x), \qquad x(0) = x_0, \tag{1.8}$$

where $x \in \mathbb{R}$ and $\alpha > 0$ is a parameter. This differential equation has equilibrium solutions $x_1^* = 0$ (unstable) and $x_2^* = 1$ (stable). The ω-limit set of all solutions with $x_0 > 0$ is x_2^*.

The Explicit Euler Method, applied to the logistic equation, gives the map

$$x \mapsto x + h\alpha x(1-x). \tag{1.9}$$

Figure 1.1 depicts the bifurcation diagram for (1.9). As the parameter $l = h\alpha$ is increased through the value 2, x_2^* loses stability and the system undergoes a series of spurious period doubling bifurcations.

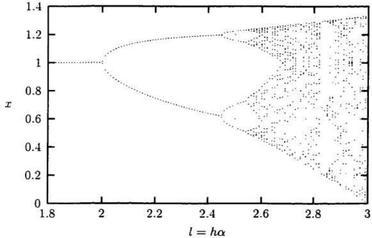

Figure 1.1: Bifurcation diagram: Explicit Euler method, applied to (1.8).

In the simple example above, no spurious behaviour exists below the linearized stability limit ($l = 2$) of the method for the fixed point x_2^*. However, spurious behaviour *has* been observed by Yee, Sweby and their collaborators (Yee *et al.* [35]) below linearized stability limits of methods.

1.3 Invariant Sets and Manifolds

Definition 1.3.1 *A set $S \subset \mathbb{R}^m$ is said to be invariant under the vector field (1.2) if for any $x_0 \in S$, the solution $x(t, x_0) \in S$ for all $t \in \mathbb{R}$.*

Let x^* be an equilibrium point of (1.2). In seeking to study the nature of the orbits of the system near x^*, we consider the associated linear equation

$$x' = f_x(x^*)x. \tag{1.10}$$

The following sets (or manifolds) are invariant under this linear flow.

$$\begin{aligned} E^s &= \text{span}\{e_1, e_2, \cdots, e_s\} \\ E^u &= \text{span}\{e_{s+1}, e_{s+2}, \cdots, e_{s+u}\} \\ E^c &= \text{span}\{e_{s+u+1}, e_{s+u+2}, \cdots, e_{s+u+c}\}, \end{aligned}$$

where $s + u + c = m$ and $e_1, \cdots, e_s$ are the eigenvectors of $f_x(x^*)$ corresponding to the eigenvalues of $f_x(x^*)$ having negative real part, $e_{s+1}, \cdots, e_{s+u}$ are the eigenvectors of $f_x(x^*)$ corresponding to the eigenvalues of $f_x(x^*)$ having positive real part, and $e_{s+u+1}, \cdots, e_{s+u+c}$ are the eigenvectors of $f_x(x^*)$ corresponding to the eigenvalues of $f_x(x^*)$ having zero real part. The sets E^s, E^u and E^c are invariant subspaces of $\mathbb{R}^m$, and are referred to as stable, unstable and center subspaces of the linearization (1.10). Furthermore, the subspaces E^s and E^u have the following asymptotic properties:

(i) Solutions with initial values in E^s approach x^* asymptotically as $t \to +\infty$.

(ii) Solutions with initial values in E^u approach x^* asymptotically as $t \to -\infty$.

Considering the nonlinear system (1.2), we define the global stable and unstable manifolds of x^* respectively as

$$\begin{aligned} W^s(x^*) &= \{x_0 \in \mathbb{R}^m | \Phi(x_0, t) \to x^* \text{ as } t \to +\infty\}, \\ W^u(x^*) &= \{x_0 \in \mathbb{R}^m | \Phi(x_0, t) \to x^* \text{ as } t \to -\infty\}. \end{aligned}$$

These manifolds are tangent to the respective invariant subspaces of the linear vector field (1.10) at x^*; hence they are locally representable as graphs.

To discuss center manifolds, we assume without loss of generality that $x^* = 0$ and (1.2) has been transformed to the form

$$\begin{aligned} x_1' &= Ax_1 + u(x_1, x_2) \\ x_2' &= Bx_2 + v(x_1, x_2), \qquad x = (x_1, x_2) \in \mathbb{R}^c \times \mathbb{R}^s, \end{aligned} \tag{1.11}$$

where $u(0,0) = 0$, $v(0,0) = 0$, $u_{x_1}(0,0) = 0$ and $v_{x_2}(0,0) = 0$. In (1.11), A is a $c \times c$ matrix having eigenvalues with zero real parts, B is an $s \times s$ matrix having eigenvalues with negative real parts, and u and v are C^r functions ($r \geq 2$). Then, an invariant manifold that is locally representable as follows

$$W^c(0) = \{(x_1, x_2) \in \mathbb{R}^c \times \mathbb{R}^s | x_2 = h(x_1), |x_1| < \delta, h(0) = 0, h'(0) = 0\}$$

for sufficiently small δ, is called a *center manifold* for (1.11). A center manifold, $W^c(0)$ is tangent to E^c at $x = (x_1, x_2) = (0, 0)$.

The *Center Manifold Theorem* (see Guckenheimer & Holmes [11], Wiggins [32]) states that the dynamics of (1.11) restricted to the center manifold is given by the vector field

$$y' = Ay + u(y, h(y)), \qquad y \in \mathbb{R}^c,$$

and, near $y = 0$, determines the dynamics of (1.11) near $(x_1, x_2) = (0, 0)$.

Analogous definitions for maps exist.

Definition 1.3.2 *A set $S \subset \mathbb{R}^m$ is said to be invariant under the map (1.3) if for any $x_0 \in S$, $g^n(x_0) \in S$ for all n.*

Let X^* be a fixed point of (1.3). In seeking to study the nature of the orbits of the system near X^*, we consider the associated linear map

$$x \mapsto g_x(X^*)x. \qquad (1.12)$$

The following sets (or manifolds) are invariant under this linear map.

$$E^s = \text{span}\{e_1, e_2, \cdots, e_s\} \tag{1.13}$$

$$E^u = \text{span}\{e_{s+1}, e_{s+2}, \cdots, e_{s+u}\} \tag{1.14}$$

$$E^c = \text{span}\{e_{s+u+1}, e_{s+u+2}, \cdots, e_{s+u+c}\} \tag{1.15}$$

where $s+u+c = m$ and $e_1, \cdots, e_s$ are the eigenvectors of $g_x(X^*)$ corresponding to the eigenvalues of $g_x(X^*)$ having modulus less than one, $e_{s+1}, \cdots, e_{s+u}$ are the eigenvectors of $g_x(X^*)$ corresponding to the eigenvalues of $g_x(X^*)$ having modulus greater than one, and $e_{s+u+1}, \cdots, e_{s+u+c}$ are the eigenvectors of $g_x(X^*)$ corresponding to the eigenvalues of $g_x(X^*)$ having modulus one. E^s, E^u and E^c are invariant subspaces of $\mathbb{R}^m$, and are referred to as stable, unstable and center subspaces. The subspaces E^s and E^u have the following asymptotic properties:

(i) Solutions with initial values in E^s approach X^* asymptotically as $n \to +\infty$;

(ii) Solutions with initial values in E^u approach X^* asymptotically as $n \to -\infty$.

Considering the nonlinear map (1.3), we define the global stable and unstable manifolds of X^* respectively as

$$\begin{aligned} W^s(X^*) &= \{x_0 \in \mathbb{R}^m | \phi(t, x_0) \to X^* \text{ as } t \to +\infty\}, \\ W^u(X^*) &= \{x_0 \in \mathbb{R}^m | \phi(t, x_0) \to X^* \text{ as } t \to -\infty\}. \end{aligned}$$

These manifolds are tangent to the respective invariant subspaces of the linear vector field (1.12) at X^*.

The discussion of center manifolds for maps is very similar to the one for vector fields; we assume without loss of generality that $X^* = 0$ and (1.3) has been transformed to the following form

$$\begin{aligned} x_1 &\mapsto Ax_1 + u(x_1, x_2) \\ x_2 &\mapsto Bx_2 + v(x_1, x_2), \qquad x = (x_1, x_2) \in \mathbb{R}^c \times \mathbb{R}^s, \end{aligned} \tag{1.16}$$

where $u(0,0) = 0$, $v(0,0) = 0$, $u_{x_1}(0,0) = 0$ and $v_{x_2}(0,0) = 0$. In (1.16), as in (1.11), A is a $c \times c$ matrix having eigenvalues of modulus one, B is an $s \times s$ matrix having eigenvalues of modulus less than one, and u and v are C^r functions ($r \geq 2$). Then there exists an invariant manifold that is locally representable as follows

$$W^c(0) = \{(x_1, x_2) \in \mathbb{R}^c \times \mathbb{R}^s | x_2 = h(x_1), |x_1| < \delta, h(0) = 0, h'(0) = 0\}$$

for sufficiently small δ. This manifold is called a *center manifold* for (1.16). A center manifold, $W^c(0)$ is tangent to E^c at $x = (x_1, x_2) = (0, 0)$.

The Center Manifold Theorem for maps is completely analogous to the one given for vector fields and states that the dynamics of (1.16) restricted to the center manifold is given by the map

$$y \mapsto Ay + u(y, h(y)), \qquad\qquad y \in \mathbb{R}^c,$$

and, near $y = 0$, determines the dynamics of (1.16) near $(x_1, x_2) = (0, 0)$.

1.4 Nonautonomous Case

Consider the ODE

$$x' = f(t, x), \qquad x(0) = x_0, \tag{1.17}$$

where $f : I \subset \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, and f depends explicitly on t.

We confine ourselves to the case in which f is assumed to be periodic in t with prime period T. Under certain conditions on f, (1.17) has a unique T-periodic solution. These will be stated for each form of f we will consider. In these cases, the notion of equilibrium points encountered in autonomous systems is replaced by that of periodic solutions.

Definition 1.4.1 *A periodic solution, $\Phi(t, x_0)$, of (1.17) is said to be* stable *if, for any $\epsilon > 0$, there is a $\delta = \delta(\epsilon) > 0$, such that, for every y_0 for which*

$|y_0 - x_0| < \delta$, *the solution* $\Phi(t, y_0)$ *of (1.17) satisfies* $|\Phi(t, y_0) - \Phi(t, x_0)| < \epsilon$ *for all* $t \geq 0$.

$\Phi(t, x_0)$ *is said to be asymptotically stable if it is stable and, in addition, there is an* $r > 0$ *such that* $|\Phi(t, y_0) - \Phi(t, x_0)| \rightarrow 0$ *as* $t \rightarrow +\infty$ *for all* y_0 *satisfying* $|y_0 - x_0| < r$.

We will use linearized one-point collocation methods to study the dynamics of the nonautonomous system (1.17) for the special forms of f:

(i) f linear: $f(x, t) = a(t)x + b(t)$, where $a(t)$ and $b(t)$ are C^1 T-periodic functions of t.

(ii) f nonlinear: $f = a(t)g(x) + b(t)$, where $a(t)$ and $b(t)$ are as in (i) and $g(x)$ is a C^2 nonlinear function of x.

The *Poincaré map* is often used to study the stability of periodic solutions.

Definition 1.4.2 *Let (1.17) be a* T*-periodic scalar ordinary differential equation with* T*-periodic solution* $\Phi(t, x_0)$. *Then, the Poincaré map of (1.17) is the scalar mapping*

$$\Pi : \mathbb{R} \rightarrow \mathbb{R}; \qquad x_0 \mapsto \Phi(T, x_0).$$

A point x_0 is the initial value of a T-periodic solution of (1.17) if and only if $\Pi(x_0) = x_0$, that is, it is a fixed point of the Poincaré map. The stability properties of the T-periodic solution are the same as the stability properties of the corresponding fixed point of the Poincaré map (see Hale & Koçac [16]). We will use an analogous procedure; by selecting the stepsize h in a numerical method such that T is an integral multiple of h, $T = hk$ ($k \in \mathbb{N}$), we can use the k-fold composition of the method instead of the Poincaré map. This procedure, which does not require prior knowledge of the periodic solution, is known as *stroboscopic sampling*, and is explained in explicit detail in Chapter 5.

1.5 Linearized Collocation Methods

Our study will be based on *linearized one-point collocation methods*, which is a class of numerical methods for the solution of (1.1). We will present a short derivation of the methods, and present three important special cases.

1.5.1 Derivation of One-Point Collocation Methods

We approximate the solution, $x(t)$, by a piecewise linear $\mathbf{C}^0$ function, $u(t)$; on $[t_n, t_{n+1}]$, $u(t)$ is given by

$$u(t_n + sh) = x_n + hsz_n \tag{1.18}$$

where $s \in [0, 1]$ and $z_n = u'(t)$ is constant on (t_n, t_{n+1}).

Collocation at $t = t_n + c_1 h$, where $0 \leq c_1 \leq 1$ is given, implies

$$u'(t_n + c_1 h) = f(t_n + c_1 h, u(t_n + c_1 h)) \tag{1.19}$$

and, from (1.18) and (1.19),

$$z_n = f(t_n + c_1 h, x_n + hc_1 z_n). \tag{1.20}$$

If we let $s = 1$ in (1.18), noting that $u(t_n) = x_n$, $u(t_{n+1}) = x_{n+1}$, then

$$x_{n+1} = x_n + hz_n. \tag{1.21}$$

Rearranging (1.20) and (1.18) gives the family of numerical methods called one-point collocation methods:

$$x_{n+1} = x_n + hf(t_n + c_1 h, (1 - c_1)x_n + c_1 x_{n+1}), \tag{1.22}$$

where $c_1 \in [0, 1]$.

Collocation methods for initial value problems were introduced in Loscalzo [24], and have attracted great interest ever since. A general discussion of the methods can be found in Hairer *et al.* [14], and a brief history is given in Brunner [5]. All the one-point methods are globally first-order; however, Guillou and Soulé [12] proved that local superconvergence of the methods is attained at the mesh points if the $\{c_i\}$ are Gauss points. In the one-point case, this would mean order 2 local superconvergence is attained when $c_1 = 0.5$.

Special Cases: Note that putting $c_1 = 0$ in (1.22) gives us the explicit Euler method, $c_1 = 1/2$ gives the implicit midpoint method, and $c_1 = 1$ gives the implicit Euler method.

One-point collocation methods belong to a class of continuous one-step Runge-Kutta methods for (1.1); they are implicit whenever $c_1 > 0$.

When $c_1 > 0$, difficulties can arise (such as loss of uniqueness) in attempting to solve the implicit equations at each step for x_{n+1}. A method that has proved to be numerically inexpensive and reliable (see Yee & Sweby [33]) is linearization. Also, the study of the dynamics of linearized methods is of interest on its own; due to their obvious computational advantages, they would be preferable over fully implicit methods if they retain the dynamics of the differential equation.

1.5.2 Linearization

We perform a linearization of (1.22) about the point (t_n, x_n). The resulting methods take the following form:

$$x_{n+1} = x_n + h[I - c_1 h f_x(t_n, x_n)]^{-1}[f(t_n, x_n) + c_1 h f_t(t_n, x_n)]. \tag{1.23}$$

Extracting the three special cases gives:

Explicit Euler ($c_1 = 0$):

$$x_{n+1} = x_n + hf(t_n, x_n). \tag{1.24}$$

Linearized Implicit Midpoint ($c_1 = 0.5$):

$$x_{n+1} = x_n + h[I - \frac{h}{2}f_x(t_n, x_n)]^{-1}[f(t_n, x_n) + \frac{h}{2}f_t(t_n, x_n)]. \tag{1.25}$$

Linearized Implicit Euler ($c_1 = 1$):

$$x_{n+1} = x_n + h[I - hf_x(t_n, x_n)]^{-1}[f(t_n, x_n) + hf_t(t_n, x_n)]. \tag{1.26}$$

Observe that for autonomous equations, linearized one-point collocation methods are given by

$$x_{n+1} = x_n + h[I - c_1 h f_x(x_n)]^{-1} f(x_n). \tag{1.27}$$

The following lemmas are simple consequences of the form of the linearized one-point collocation methods (1.27).

Lemma 1.5.1 *If $x = x^*$ is an equilibrium point of $x' = f(x)$, and $hc_1 f_x(x^*) \neq I$, then x^* is a fixed point of (1.27).*

Lemma 1.5.2 *If $x = X^*$ is a fixed point of (1.27), then it is an equilibrium point of the generating vector field $x' = f(x)$ of (1.27).*

Proof: We write (1.27) in the form

$$[I - c_1 h f_x(x_n)](x_{n+1} - x_n) = h f(x_n).$$

Clearly, if X^* is a fixed point of (1.27), $f(X^*) = 0$.

1.6 Outline of Thesis

Chapter 2 gives some background and motivation of the work in the thesis, and provides some historical development of the area of Dynamics of Numerics.

In Chapter 3, we apply the linearized one-point collocation methods to autonomous equations of the form $x' = f(\mu, x)$, where μ is a parameter. Our objective is to ascertain whether any bifurcations that occur in the differential equation are inherited by the methods, and at the same values of the parameter.

In Chapter 4, we consider the possible effect of the singularity in these collocation methods when solving autonomous equations. It is shown that the singularity structures of the methods are responsible for the distortion of the global asymptotic behaviour of the methods.

In Chapter 5, we consider nonautonomous equations. Starting with linear equations, we establish conditions under which the numerical methods' dynamical behaviour mimics that of the differential equations. This study is extended to nonlinear problems, and it is shown that there is a relationship between dynamical behaviour and linear/nonlinear stability theory for our chosen methods. The study is extended to multi-dimensional linear nonautonomous equations in Chapter 6.

Chapter 7 summarizes the results obtained in the thesis, and gives some suggestions for future work.

Chapter 2

History and Background

2.1 Introduction

The subject treated in this thesis belongs to a new area of Applied Mathematics called "The Dynamics of Numerics", which attempts to study numerical methods for the solution of differential equations using the theory of dynamical systems.

The study of numerical schemes using the theory of dynamical systems has attracted much interest since the early 1990's and the term "Dynamics of Numerics", as applied in this context, was the subject of the Conference on Dynamics of Numerics and Numerics of Dynamics [4].

It is now well established that numerical methods are dynamical systems in their own right whose behaviour can differ significantly, often drastically, from that of the differential equation they are attempting to solve. The implications of this possibility are far reaching. In practice, to predict long term behaviour of solutions of differential equations, one has to rely solely on numerical schemes. Therefore, it is absolutely essential to be familiar with the dynamical behaviour of the numerical scheme one intends to use. The importance of this investigation is stated explicitly by Stewart [29]:

> "... the safest route is to have some understanding of the dynamical behaviour of the numerical method being used. In short, the whole area needs sorting out. The main problems are not so much numerical as dynamical: the actual behaviour of the continuum models, and their relation to discretizations, must be the central object of study. Until these advances in the dynamics of numerics are made, all users of the numerics of dynamics — most of whom are wielding a mathematical tool without understanding how it works — should heed the warning."

This chapter will survey existing results on this dynamic subject which has come to be known as the Dynamics of Numerics. The text by Stuart & Humphries [31] gives a comprehensive review of the then known results on this subject.

We consider the autonomous differential equation

$$x' = f(x), \qquad x(0) = x^0 \in \mathbb{R}^m,$$

and the corresponding map associated with a numerical discretization

$$x \mapsto x + h\phi(x,h) := g(x,h).$$

It is the asymptotic states (the α- and ω-limit sets) that are of interest in this chapter.

2.2 Spurious Fixed Points

Iserles [19] made the connection between the dynamical features of the map representing a numerical method and those of the autonomous differential equation (1.2). Defining the sets

$$\begin{aligned} F &:= \{x \in \mathbb{R}^m : f(x) = 0\} \qquad \text{and} \\ G_h &:= \{X \in \mathbb{R}^m : \phi(X,h) = 0\}, \end{aligned}$$

he proved the following results.

Theorem 2.2.1 *For linear multi-step methods, $G_h = F$ for all $h > 0$.*

Theorem 2.2.2 *For Runge-Kutta and Predictor-Corrector methods, $F \subseteq G_h$, and the inclusion may be strict.*

In other words, linear multi-step methods are $R^{[1]}$ whereas Runge-Kutta and Predictor-Corrector methods are generally $IR^{[1]}$ (can admit spurious fixed points). However, even in the absence of spurious fixed points, the dynamics of a linear multi-step method can differ from that of the differential equation since the domain of attraction of a fixed point of the method may fail to resemble that of the corresponding equilibrium point in the differential equation, depending on h (see Iserles [19]).

2.3 Characterization of Regularity in Runge-Kutta Methods

Hairer *et al.* [13] studied conditions under which Runge-Kutta methods are $R^{[1]}$. They established the following results.

Theorem 2.3.1 *A consistent explicit Runge-Kutta scheme is $R^{[1]}$ if and only if it produces the same solution sequence as the explicit Euler method.*

This theorem essentially means that spurious fixed points can be expected in higher order explicit Runge-Kutta methods. The following theorem, also in Hairer *et al.* [13], states the regularity condition of general Runge-Kutta methods.

Theorem 2.3.2 *The maximal order of an $R^{[1]}$ Runge-Kutta method is 3,* *and the maximal order of an A-stable Runge-Kutta method is 4.*

The above results do not specify the magnitude of the stepsize for which spurious fixed points are observable. Hence, it may be assumed that they would occur for stepsizes above those normally used in practice. Humphries [18] studied the behaviour of these spurious fixed points as $h \to 0$, and Griffiths *et al.* [10] studied the occurrence of spurious fixed points below the linearized stability limit for the genuine fixed points.

2.4 Study of Spurious Fixed Points

Humphries [18] studied the behaviour of spurious fixed points, whose existence was uncovered by Iserles [19], in the limit as $h \to 0$. He proved the following results.

Theorem 2.4.1 *If a numerical approximation to (1.2) is obtained by using an explicit Runge-Kutta method or an implicit Runge-Kutta method (where the implicit equations are solved using a convergent iterative scheme), then*

(i) *if f is globally Lipschitz, spurious fixed points cannot exist for h arbitrarily small,*

(ii) *if f is locally Lipschitz, and in particular if $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$, and* *spurious fixed points exist for h arbitrarily small, then they tend to infinity in norm, as $h \to 0$.*

A similar result is true for period-2 solutions.

Theorem 2.4.2 *If a numerical approximation to (1.2) is obtained using an explicit R-K method, an implicit R-K method (where the implicit equations are solved using a convergent iteration), or a zero-stable linear multi-step method with $\rho(-1) \neq 0$, then*

(i) *if f is globally Lipschitz, period-2 solutions cannot exist for h arbitrarily small,*

(ii) *if f is locally Lipschitz, and in particular if $f \in C^1(\mathbb{R}^m, \mathbb{R}^m)$ and a period-2 solution $(u(h), v(h))$ exists for h arbitrarily small, then both $u(h)$, $v(h)$ tend to infinity in norm, as $h \to 0$.*

While these results are quite revealing, in practical computations, h remains finite but not extremely small, so that $h \not\to 0$. This necessitates the study of the possible effect of spurious solutions which may coexist with the true asymptotic solutions.

More recently, Stein [28] proved that, for any hyperbolic fixed point of an explicit Runge-Kutta method, there exists a computable limit stepsize h_l, below which spurious solutions do not exist and the method inherits only the fixed points of the differential equation with their stability types. While this result reveals that spurious solutions do not exist provided the stepsize is small enough, we often have to use relatively large stepsizes when computing over long time intervals.

2.5 Bifurcations to Spurious Period-2 Solutions

Iserles *et al.* [20] proved that spurious period-2 solutions can bifurcate from genuine fixed points as h is varied. This, they showed, can occur in both Runge-Kutta and linear multi-step methods. They also established an order condition for $R^{[1,2]}$ Runge-Kutta methods, which is stated below.

Theorem 2.5.1 *The highest attainable order of an $R^{[1,2]}$ Runge-Kutta method is 2.*

A natural question to ask is: How do period-2 solutions in $IR^{[2]}$ methods arise? The two theorems below answer this question for Runge-Kutta and linear multi-step methods.

Theorem 2.5.2 *Consider the Runge-Kutta method (1.5), used to solve the scalar equation $x' = f(x)$ where $f : \mathbb{R} \to \mathbb{R}$, which has a hyperbolic equilib-*

rium point x^*. Then, period-2 solutions bifurcate from x^* at $h = h_c$, where

$$\mathbf{b}^T(I - h_c f'(x^*)A)^{-1}\mathbf{1} + \frac{2}{h_c f'(x^*)} = 0, \tag{2.1}$$

provided $(I - h_c f'(x^*)A)$ is invertible and $\mathbf{b}^T(I - h_c f'(x^*)A)^{-2}\mathbf{1} \neq 0$.

A similar result holds for linear multi-step methods.

Theorem 2.5.3 *Consider the consistent and zero-stable linear multi-step method (1.4), with $\rho(-1)$, $\sigma(-1) \neq 0$. Suppose this method is used to solve the scalar equation $x' = f(x)$, where $f : \mathbb{R} \to \mathbb{R}$ which has a hyperbolic equilibrium point x^*. Then, period-2 solutions bifurcate from x^* at $h = h_c$, where*

$$h_c = \frac{\rho(-1)}{f'(x^*)\rho(-1)}. \tag{2.2}$$

Griffiths *et al.* [10] show that, even though h_c may be above the linearized stability limit of the respective method for x^*, an unstable branch of the spurious solution may exist below the linearized stability limit. This may affect the global asymptotics of the method for reasonable values of h by distorting or segmenting the basin of attraction of x^* (see Yee & Sweby [36] and Chapter 4 of the thesis).

2.6 Branches of Spurious Solutions

In [10], Griffiths *et al.* study a genuine fixed point, x^*, of an explicit Runge-Kutta method such as (1.5). Perturbation arguments are used to investigate the local nature of bifurcations from the fixed point to spurious solutions.

Linearization about a fixed point leads to the stability function

$$R(x^*, h) = 1 + \rho \mathbf{b}^T (I - \rho A)^{-1} \mathbf{1} \tag{2.3}$$

where $\rho = hf'(x^*)$, $\mathbf{b} = (b_1, b_2, \dots, b_s)^T$, $\mathbf{1} = (1, 1, \dots, 1)^T$, and A denotes the $s \times s$ array of weights a_{ij} in (1.6).

Loss of stability of x^* is encountered when $\rho = \rho^* = h^* f'(x^*)$ satisfies either

$$\rho^* \mathbf{b}^T (I - \rho^* A)^{-1} \mathbf{1} = -2 \tag{2.4}$$

or

$$\rho^* \mathbf{b}^T (I - \rho^* A)^{-1} \mathbf{1} = 0. \tag{2.5}$$

In the former case, as ρ decreases beyond ρ^*, the method undergoes a period doubling bifurcation. In the latter case, x^* loses stability and for a linear problem, this would lead to $|X_n| \to \infty$ as $n \to \infty$; however, for genuinely non-linear differential equations ($f'(x) \neq$ constant), there may be a bifurcation to a fixed point X^* that is spurious.

It is further proved that not only can spurious steady states be reached for $h > h^*$ (corresponding to ρ^*), but an unstable branch exists for $h < h^*$, which may substantially affect the domain of attraction of the true, stable steady state.

This work is enlightening because it shows that the existence of spurious solutions of a Runge-Kutta scheme, regardless of their stability, below the linearized stability limit of the scheme for a genuine fixed point, can adversely affect the computed solution and result in a distorted numerical basin of attraction.

2.7 Global Asymptotic Behaviour

Yee & Sweby ([36], [33], [34]) concentrate their study on specific 2×2 systems of ODEs and show how spurious asymptotes, regardless of their stability, can give rise to *numerical basins of attraction* that differ significantly from the basins of the ODE for the true asymptotes.

Nine explicit and two implicit R-K methods, as well as four linear multistep methods are considered. These are the explicit Euler, modified Euler, improved Euler, Heun, R-K 3, R-K 4, PC2, Adam-Bashforth, linearized implicit Euler, linearized trapezoidal, implicit Euler, trapezoidal, 3-level BDF and mid-point implicit method. Comprehensive bifurcation diagrams are obtained. All eleven R-K methods generate spurious asymptotes, and for the predator-prey model, for example, more than one spurious fixed point below the linearized stability limit of the scheme is introduced.

Numerical results indicate that, for different finite discretization parameters h_1 and h_2 below the linearized stability limit of the scheme, numerical solutions might converge to two different solutions even if no spurious stable steady-state numerical solution is introduced by the scheme. The existence of spurious asymptotes, regardless of their stability has detrimental effects on the computed solution.

Thus, for a given h below the linearized stability limit of the scheme, the numerical solution may:

(a) converge to the correct steady state;

(b) converge to a spurious steady state;

(c) converge to a spurious periodic solution;

(d) yield spurious asymptotes other than (b) or (c);

(e) diverge, even for physically relevant initial data.

Even though linear multi-step methods preserve the same number of steady states as the underlying ODE, they may change stability types of the fixed points. Also, the solution procedure used in computing the full discretized equations, has an effect on the asymptotic behaviour. Results show that linearized implicit methods are more efficient than other solution procedures.

2.8 Relevance of Thesis

This thesis will settle some issues in the dynamics of numerics that are either incompletely considered or not treated in the literature.

Collocation methods have grown in importance since their initial introduction. Generally speaking, implicit methods tend to be computationally expensive in practice, and explicit methods are computationally cheaper yet impose stringent stability restrictions which place major limits on the discretization parameters. In order to determine whether linearization gives us the best of both worlds, we have to study the dynamics of the linearized methods.

Much of the study that is represented in the literature has concentrated on autonomous ODEs, fixed points and possible existence of spurious fixed points in Runge-Kutta and linear multi-step methods. This thesis goes further, and studies the dynamics of linearized collocation methods for both autonomous equations and nonautonomous equations with periodic solutions.

From the numerical results in Yee & Sweby ([33], [36]), it is observed that there are different numerical basins of attraction for different numerical methods and different solution procedures used in implicit methods. In particular, there is a shrinking of the basin of attraction for ∞ in certain methods, particularly linearized ones. Here, we introduce the concept of pole-type behaviour and prove that the existence of spurious pole-type behaviour, which is inextricably linked to the presence of singularities in the methods, causes basin shrinkage for ∞.

It has been established in Griffiths *et al.* [10] that numerical methods can introduce spurious period-doubling bifurcations and bifurcations to spurious fixed points. What is the effect of discretization of a parameter-dependent ODE by a numerical method? In particular, we will consider two types of parameter-dependent ODEs: scalar ODEs with the common codimension-1 bifurcations and planar ODEs with the Hopf bifurcation. The objective is to ascertain whether discretization of the ODEs by linearized one-point collocation methods results in a discrete system that possesses the discrete analog of the same bifurcation and if so, whether this bifurcation occurs at the same parameter value. The results are generally applicable to higher dimensional systems through center manifold theory (as discussed in Wiggins [32]).

Chapter 3

Autonomous Equations

3.1 Introduction

Let

$$x' = f(x, \mu), \qquad \text{for } t \geq 0 \text{ and } x(0) = x^0, \tag{3.1}$$

where $f : \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$, be an ordinary differential equation with parameter $\mu \in \mathbb{R}$. Since f does not depend explicitly on t, (3.1) is *autonomous.*

In those instances in which the solution $x(t)$ of (3.1) is approximated by a linearized one-point collocation method, we would like to investigate the dynamical behaviour of the numerical method and see how far it correctly represents the dynamical features of the vector field.

The linearized one-point collocation methods, when applied to (3.1) are given by

$$X_{n+1} = X_n + h[I - hc_1 f_x(X_n, \mu)]^{-1} f(X_n, \mu) := g(X_n), \tag{3.2}$$

where I is the $m \times m$ identity matrix, f_x the Jacobian of f with respect to x and h is the time step. It is clear from Lemma 1.5.1 that $(\overline{\mu}, x^*(\overline{\mu}))$ is a fixed point of (3.2) if and only if it is an equilibrium point of (3.1) for $\mu = \overline{\mu}$ provided $hc_1 f_x(x^*(\overline{\mu}), \overline{\mu}) \neq I$.

If an equilibrium point of the system (3.1) undergoes a bifurcation, we would like to ascertain whether the fixed point in the numerical methods will undergo the discrete analog of the same bifurcation, and whether that bifurcation occurs at the same value of the parameter μ. We begin by studying common codimension-1 bifurcations, then the Hopf bifurcation in planar systems.

3.2 Codimension-1 Bifurcations

The stability of common codimension-1 bifurcations under linearized one-point collocation methods has been studied using normal forms, and detailed results are in Foster & Khumalo [7].

Normal forms is a method of simplifying a dynamical system (continuous or discrete) by finding a coordinate system in which the dynamical system takes the simplest form. The coordinate transformations generated are typically local to a known solution (such as a fixed point or periodic solution).

3.2.1 Saddle-Node Bifurcation

The following lemmas give well-known topological normal forms for one-dimensional systems that possess the saddle-node bifurcation (see Wiggins [32]).

Lemma 3.2.1 *Let $x' = f(x, \alpha)$ be a one-parameter family of scalar flows such that f is C^2 in x and C^1 in α on some sufficiently large open set containing $(0,0)$. Let the origin be a nonhyperbolic fixed point: $f(0,0) = 0$, $f_x(0,0) = 0$. Let the non-degeneracy conditions $f_{xx}(0,0) \neq 0$, $f_\alpha(0,0) \neq 0$ exist. Then a normal form for the saddle-node bifurcation is*

$$x' = \alpha \pm x^2. \tag{3.3}$$

The normal form equation $x' = \alpha + x^2$ has equilibrium points $x^* = \pm\sqrt{-\alpha}$. For $\alpha < 0$, $x^* = +\sqrt{-\alpha}$ is unstable and $x^* = -\sqrt{-\alpha}$ is stable. At $\alpha = 0$, there is a single unstable equilibrium point $x^* = 0$.

The second lemma states a normal form for the saddle-node bifurcation in maps.

Lemma 3.2.2 *Let $x \mapsto g(x, \mu)$ be a one-parameter family of scalar maps such that g is C^2 in x and C^1 in μ. Let $g(0,0) = 0$ and $g_x(0,0) = 1$, $g_{xx}(0,0) \neq 0$ and $g_\mu(0,0) \neq 0$. Then a normal form for the saddle-node bifurcation is*

$$x \mapsto \mu + x \pm x^2. \tag{3.4}$$

The curves of fixed points in equations (3.3) and (3.4) correspond exactly in location and stability type.

Theorem 3.2.3 *If a saddle-node bifurcation occurs in $x' = f(x, \alpha)$, then a corresponding saddle-node occurs in its transformation by linearized one-point collocation.*

Proof [7]: Applying linearized one-point collocation methods to the normal form (3.3) gives rise to the maps $x \mapsto g$, where

$$g = x + \frac{h(\alpha + \bar{s}x^2)}{1 - \bar{s}hc_1x},$$

and $\bar{s} = \pm 1$.

Expanding as $g(x, \alpha) = x + \sum g_i(\alpha)x^i$, we obtain $g = h\alpha + (1 + 2\bar{s}h^2c_1\alpha)\, x + (\bar{s}h + 4h^3c_1^2\alpha)\, x^2 + O(x^3)$.

Let $\xi = x+\delta$, where $\delta = \delta(\alpha)$ is to be defined. Then the image of ξ under the action of the map is

$$\begin{aligned}\xi &= x+\delta \mapsto g(x,\alpha)+\delta = g(\xi-\delta,\alpha)+\delta \\ &= [g_0 - g_1\delta + g_2\delta^2 + O(\delta^3)] + [1+g_1-2g_2\delta+O(\delta^2)]\xi + [g_2+O(\delta)]\xi^2 + O(\xi^3).\end{aligned}$$

Because $g_1(0) = 0$ and $g_2(0) \neq 0$, the Implicit Function Theorem can be invoked to annihilate the α-dependent part of the linear term (in ξ) for all sufficiently small $|\alpha|$. Hence we define $\delta(\alpha) = \dfrac{g_1'(0)}{2g_2(0)}\alpha + O(\alpha^2) = hc_1\alpha + O(\alpha^2)$, obtaining

$$\xi \mapsto [h\alpha - \bar{s}h^3c_1^2\alpha^2 + \alpha^3\Phi(\alpha)] + \xi + [\bar{s}h + O(\alpha)]\,\xi^2 + O(\xi^3), \tag{3.5}$$

where $\Phi(\alpha)$ is a smooth function. Let $\gamma(\alpha)$ be the ξ-independent term of (3.5). Since $\gamma(0) = 0$ and $\gamma'(0) = h > 0$, the Inverse Function Theorem assures local existence and uniqueness of a smooth inverse $\alpha(\gamma)$ increasing through the origin. Then $\xi \mapsto \gamma+\xi+a_2(\gamma)\xi^2+O(\xi^3)$, where $a_2(\gamma)$ is smooth with $a_2(0) = \bar{s}h \neq 0$. A final change of variable $\eta = |a_2(\gamma)|\,\xi$ gives

$$\eta \mapsto \mu + \eta + \bar{s}\eta^2 + O(\eta^3),$$

where $\mu = |a_2(\gamma)|\,\gamma$. The higher-degree terms can be eliminated due to local topological conjugacy of these continuously parameter-dependent maps [2].

Thus we obtain the normal form (3.4) for saddle-node bifurcation at $(\eta, \mu) = (0, 0)$. Direct consideration of the above transformation sequence, restricted to a neighbourhood of $(x, \alpha) = (0, 0)$, confirms that $(\eta, \mu) = (0, 0)$ if and only if $(x, \alpha) = (0, 0)$. Furthermore, the derived map and the original flow have exactly corresponding normal forms and μ is x-independent. Therefore, the phase location, orientation and stability properties of the saddle-node bifurcation are preserved under the map. □

We now consider the stability of fixed points. Assuming without loss of generality that $\tilde{s} = +1$, the fixed points are given by $x^* = \pm\sqrt{-\alpha}$. Singularities occur for $1 - 2hc_1x = 0$, hence the fixed points exist for $\alpha \le 0$, $\alpha \ne -1/(4h^2c_1^2)$. Performing standard stability analysis, we determine that the fixed point $x^* = \sqrt{-\alpha}$ is stable for

$$\left\{ \alpha \overset{PD}{<} \frac{-1}{h^2\,(2c_1 - 1)^2} \right. \qquad (c_1 > 1/2).$$

and the fixed point $x^* = -\sqrt{-\alpha}$ is stable for

$$\begin{cases} \alpha \overset{SN}{<} 0 & (c_1 \ge 1/2) \\ \dfrac{-1}{h^2\,(1 - 2c_1)^2} \overset{PD}{<} \alpha \overset{SN}{<} 0 & (c_1 < 1/2). \end{cases}$$

The labelled inequalities represent the locations of saddle-node (SN) and period doubling (PD) bifurcations, see Figures 3.1 and 3.2. The following key applies to each of the diagrams:

Stable fixed point: —

Unstable fixed point: – – –

Singular set: - - - -

Observe that when $c_1 = 0$ (explicit Euler method), the singularity is removed to $x \to \infty$. On the other hand, for $c_1 = 1/2$ (linearized implicit midpoint) the period doubling bifurcation is removed to $x \to \infty$.

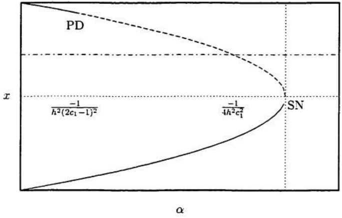

Figure 3.1: Saddle-node: $f = \alpha + x^2$, $c_1 > 0.5$

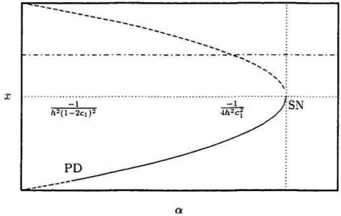

Figure 3.2: Saddle-node: $f = \alpha + x^2$, $c_1 < 0.5$

We now consider the converse of Theorem 3.2.3. That is, we assume a saddle-node bifurcation exists in the family of maps generated by linearized one-point collocation methods, and examine the nature of the flow from which the maps are derived.

Theorem 3.2.4 *If a saddle-node bifurcation occurs in a scalar map $x \mapsto g(x, \mu)$ due to a transformation by a linearized one-point collocation method, it results from a corresponding saddle-node in the originating scalar flow $x' = f(x, \alpha)$.*

Proof [7]: Assume a linearized one-point collocation method possesses a saddle-node bifurcation at $(x, \mu) = (0, 0)$, where $f = f(x, \mu)$ is an undetermined flow function given to be C^2 in x and C^1 in μ. Equating this family of maps to the normal form (3.4), and expanding $f(x, \mu) = \sum a_i(\mu)x^i$, we match coefficients to obtain

$$f(x, \mu) = \frac{\mu}{h} + 2c_1^2\mu^2 a_2(\mu) - 2c_1\mu a_2(\mu)x + a_2(\mu)x^2 + O(x^3).$$

A change of variable $\xi = x + \delta(\mu)$ yields the originating flow as

$$\begin{aligned} \xi' &= x' = f(x, \mu) = f(\xi - \delta, \mu) \\ &= [a_0 - a_1\delta + a_2\delta^2 + O(\delta^3)] + [a_1 - 2a_2\delta + \delta^2\Psi(\mu, \delta)]\xi + [a_2 + O(\delta)]\xi^2 + O(\xi^3), \end{aligned}$$

where $\Psi(\mu, \delta)$ is a smooth function. If $a_1(0) = 0$ and $-2a_2(0) \neq 0$, then the linear term in ξ can be annihilated for all sufficiently small $|\mu|$ by application of the Implicit Function Theorem. These conditions are satisfied whenever $\beta_0 \neq 0$ in $a_2(\mu) = \beta_0 + O(\mu)$, equivalent to the non-degeneracy condition $f_{xx}(0, 0) \neq 0$ for saddle-node bifurcation. Assuming this, we can define $\delta(\mu) = \frac{a_1'(0)}{2a_2(0)}\mu + O(\mu^2) = -c_1\mu + O(\mu^2)$, to obtain

$$f(x, \mu) = \left[\frac{\mu}{h} + \beta_0 c_1^2\mu^2 + \mu^3\Phi(\mu)\right] + [\beta_0 + O(\mu)]\,\xi^2 + O(\xi^3), \tag{3.6}$$

where $\Phi(\mu)$ is smooth.

Now define $\gamma = \gamma(\mu)$ as the ξ-independent term of (3.6). Since $\gamma(0) = 0$ and $\gamma'(0) \neq 0$, the Inverse Function Theorem guarantees local existence and uniqueness of a smooth inverse $\mu = \mu(\gamma)$ with $\mu(0) = 0$. Therefore, $\xi' = \gamma + a_2(\gamma)\xi^2 + O(\xi^3)$, with $a_2(\gamma)$ a smooth function such that $a_2(0) = \beta_0$.

Finally, we scale the variable as $\eta = |\beta_0|\,\xi$, which gives

$$\eta' = \alpha + s\eta^2 + O(\eta^3),$$

where $\alpha = |\beta_0|\left(\frac{\mu}{h} + \beta_0 c_1^2 \mu^2 + \mu^3 \Phi(\mu)\right)$ and $s = \mathrm{sgn}(\beta_0)$.

The higher-degree terms for η can be eliminated due to local topological equivalence of the flows, to attain the normal form (3.3) in η. The origin is locally preserved under the transformation $(x, \mu) \to (\eta, \alpha)$, so the location of the bifurcation is preserved. Lemma 1.5.2 implies the direction of saddle-node bifurcation is preserved under the transformation (*n.b.*, direction is not preserved by the normal forms themselves). The derived normal form is identical to (3.3), and α increases with increasing μ near the origin, so the stability type is preserved as well. □

Thus, no spurious saddle-node bifurcations can occur in linearized one-point collocation methods.

3.2.2 Transcritical Bifurcation

Intuitively, a transcritical bifurcation occurs when two equilibrium (or fixed) points (one stable, the other unstable) coalesce at the bifurcation point and an exchange of stability occurs. We now state well-known topological normal forms for the transcritical bifurcation in flows and maps.

Lemma 3.2.5 *Let $x' = f(x, \alpha)$ be a one-parameter family of scalar flows such that f is C^2 in x and C^1 in α. Assume the genericity conditions $f(0,0) = 0$, $f_x(0,0) = 0$, $f_\alpha(0,0) = 0$, and the non-degeneracy conditions $f_{xx}(0,0) \neq 0$, $f_{x\alpha}(0,0) \neq 0$. Then a normal form for the transcritical bifurcation is*

$$x' = \alpha x \pm x^2. \tag{3.7}$$

The following lemma contains a normal form for the transcritical bifurcation in a discrete system.

Lemma 3.2.6 *Let $x \mapsto g(x, \mu)$ be a one-parameter family of scalar maps such that g is C^2 in x and C^1 in μ. Assume $g(0,0) = 0$, $g_x(0,0) = 1$, $g_\mu(0,0) = 0$, $g_{xx}(0,0) \neq 0$ and $g_{x\mu}(0,0) \neq 0$. Then a normal form for the transcritical bifurcation is*

$$x \mapsto (1 + \mu)x \pm x^2. \tag{3.8}$$

The theorem below guarantees the existence of a corresponding transcritical bifurcation in the discrete system generated by the transformation of a differential equation that possesses a transcritical bifurcation.

Theorem 3.2.7 *If a transcritical bifurcation occurs in* $x' = f(x, \alpha)$, *then a corresponding transcritical bifurcation occurs in its transformation due to linearized one-point collocation methods.*

Proof: See [7].

The bifurcation behaviour of the family of maps arising from the transformation due to linearized one-point collocation can be found explicitly. The collocation methods transform (3.7) into the family of maps $x \mapsto g$, where

$$g = x + \frac{h(\alpha x + \bar{s}x^2)}{1 - hc_1(\alpha + 2\bar{s}x)}.$$

Observe that the maps introduce a singular set $1 - hc_1(\alpha + 2\bar{s}x)$, which is, however, removed from the origin.

We now obtain the bifurcation behaviour of the maps explicitly. The fixed point $x^* = 0$ is stable for

$$\begin{cases} \dfrac{-2}{h(1-2c_1)} \overset{PD}{<} \alpha \overset{T}{<} 0 & (c_1 < 1/2) \\ \alpha \overset{T}{<} 0 & (c_1 \geq 1/2) \\ \alpha \overset{PD}{>} \dfrac{2}{h(2c_1-1)} & (c_1 > 1/2). \end{cases}$$

The other fixed point is $x^* = \mp\alpha$, with the following stability regions:

$$\begin{cases} 0 \overset{T}{<} \alpha \overset{PD}{<} \dfrac{2}{h(1-2c_1)} & (c_1 < 1/2) \\ \alpha \overset{T}{>} 0 & (c_1 \geq 1/2) \\ \alpha \overset{PD}{<} \dfrac{-2}{h(2c_1-1)} & (c_1 > 1/2). \end{cases}$$

The labelled inequalities represent the locations of transcritical (T) and period doubling (PD) bifurcations. See Figures 3.3 and 3.4.

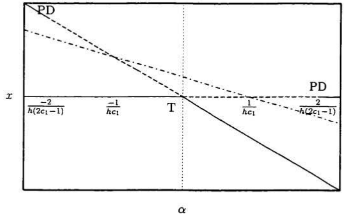

Figure 3.3: Transcritical: $f = \alpha x + x^2$, $c_1 > 0.5$

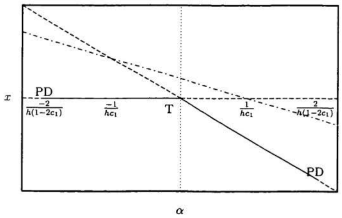

Figure 3.4: Transcritical: $f = \alpha x + x^2$, $c_1 < 0.5$

Both fixed points exist for all α, except for the singularities $(x, \alpha) = (0, (hc_1)^{-1})$, $(\pm(hc_1)^{-1}, -(hc_1)^{-1})$. $c_1 = 0$, the singularities are removed to ∞, and for $c_1 = 1/2$ the locations of the period doubling bifurcations are removed to ∞.

Theorem 3.2.8 *If a transcritical bifurcation occurs in a scalar map $x \mapsto g(x, \mu)$ due to transformation by linearized one-point collocation, it results from a corresponding transcritical bifurcation in the originating scalar flow $x' = f(x, \alpha)$.*

Proof: See [7].

Hence, linearized one-point collocation methods can introduce no spurious transcritical bifurcations.

3.2.3 Pitchfork Bifurcation

Suppose the scalar flow $x' = f(x, \alpha)$ has an equilibrium point $x^* = 0$ for all $\alpha \in \mathbb{R}$. Assume that on one side of the bifurcation value, $\alpha = 0$, the system has two other equilibrium points and on the other x^* is the only equilibrium point. Then, the system undergoes a *pitchfork* bifurcation at $\alpha = 0$. The following lemmas supply well-known topological normal forms for this type of bifurcation (Wiggins [32]).

Lemma 3.2.9 *Let $x' = f(x, \alpha)$ be a one-parameter family of scalar flows such that f is C^3 in x and C^1 in α. Assume the genericity conditions $f(0,0) = 0$, $f_x(0,0) = 0$, $f_\alpha(0,0) = 0$, $f_{xx}(0,0) = 0$, and the non-generacy conditions $f_{x\alpha}(0,0) \neq 0$, $f_{xxx}(0,0) \neq 0$. Then a normal form for the pitchfork bifurcation is*

$$x' = \alpha x \mp x^3. \tag{3.9}$$

Lemma 3.2.10 *Let $x \mapsto g(x, \mu)$ be a one-parameter family of scalar maps such that g is C^3 in x and C^1 in μ. Assume $g(0,0) = 0$, $g_x(0,0) = 1$, $g_\mu(0,0) = 0$, $g_{xx}(0,0) = 0$, $g_{x\mu}(0,0) \neq 0$ and $g_{xxx}(0,0) \neq 0$. Then a normal form for the pitchfork bifurcation is*

$$x \mapsto (1 + \mu)\, x \mp x^3. \tag{3.10}$$

The following theorem asserts that linearized one-point collocation methods do not give rise to spurious pitchfork bifurcations.

Theorem 3.2.11 *If a pitchfork bifurcation occurs in $x' = f(x, \alpha)$, then a corresponding pitchfork bifurcation occurs in its transformation due to linearized one-point collocation.*

Proof: See [7].

Explicit calculation from the family of maps arising from the transformation shows the fixed point $x^* = 0$ to be stable for

$$\begin{cases} \dfrac{-2}{h(1-2c_1)} \overset{PD}{<} \alpha \overset{P}{<} 0 & (c_1 < 1/2) \\ \alpha \overset{P}{<} 0 & (c_1 \geq 1/2) \\ \alpha \overset{PD}{>} \dfrac{2}{h(2c_1-1)} & (c_1 > 1/2). \end{cases}$$

If the pitchfork is supercritical, the nontrivial fixed points are $x^* = \pm\sqrt{\alpha}$, stable for

$$\begin{cases} 0 \overset{P}{<} \alpha \overset{PD}{<} \dfrac{1}{h(1-2c_1)} & (c_1 < 1/2) \\ \alpha \overset{P}{>} 0 & (c_1 \geq 1/2). \end{cases}$$

If the pitchfork is subcritical, the nontrivial fixed points are $x^* = \pm\sqrt{-\alpha}$, which are always unstable when $c_1 < 1/2$, and stable for

$$\begin{cases} \alpha \overset{PD}{<} \dfrac{-1}{h(2c_1-1)} & (c_1 > 1/2). \end{cases}$$

The labelled inequalities represent the locations of pitchfork (P) and period doubling (PD) bifurcations. See Figures 3.5 and 3.6.

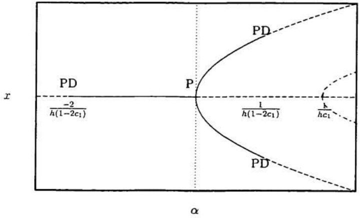

Figure 3.5: Supercritical Pitchfork: $f = \alpha x - x^3$, $c_1 < 0.5$

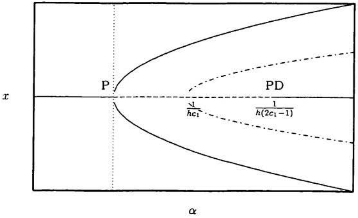

Figure 3.6: Supercritical Pitchfork: $f = \alpha x - x^3$, $c_1 > 0.5$

The nontrivial fixed points exist for $\alpha > 0$ in the supercritical case, and for $\alpha < 0$, $\alpha \neq -1/(2hc_1)$, in the subcritical case. The trivial fixed point exists for all $\alpha \neq 1/(hc_1)$. The limiting case $c_1 = 0$ merely removes the singular set to ∞. The limiting case $c_1 = 1/2$ removes the period doubling bifurcations to ∞.

Theorem 3.2.12 *If a pitchfork bifurcation occurs in a scalar map $x \mapsto g(x, \mu)$ due to transformation by a linearized one-point collocation method, it results from a corresponding pitchfork bifurcation in the originating scalar flow $x' = f(x, \alpha)$.*

Proof: See [7].

3.2.4 Period Doubling Bifurcation

The case of an eigenvalue $\lambda = -1$ is fundamentally different and does not have an analogue with one-dimensional vector field dynamics; if it occurs in a numerical method, it is necessarily spurious. This case results in a period doubling (flip) bifurcation.

Lemma 3.2.13 *Let $x \mapsto g(x, \mu)$ be a one-parameter family of scalar maps such that g is C^3 in x and C^1 in μ. Assume the genericity conditions $g(0, 0) =$*

0, $g_x(0,0) = -1$, *and the non-degeneracy conditions* $g_\mu(0,0) \neq 0$, $g^2_{xxx}(0,0) \neq 0$. *Then a normal form for the period doubling bifurcation is*

$$x \mapsto -(1+\mu)x \pm x^3. \tag{3.11}$$

The map has a nonhyperbolic fixed point with $\lambda = -1$, and the second iterate of the map undergoes a pitchfork bifurcation at the same fixed point.

Theorem 3.2.14 *Spurious period doubling bifurcations can occur under the transformation resulting from linearized collocation methods for all* $c \neq 1/2$. *A normal form for spurious period doubling bifurcation is*

$$x' = \left(\frac{2}{h(2c_1 - 1)} - \alpha\right) x \pm x^3,$$

for all $c_1 \neq 1/6,\ 1/2$.

Proof [7]: Assume the induced scalar map will have normal form (3.11) for period doubling bifurcation at $(x, \mu) = (0, 0)$. Expanding the undetermined flow function $f(x, \mu) = \sum a_i(\mu)x^i$ and solving for coefficients of powers of x in

$$-(1+\mu)x + \tilde{s}x^3 = x + hf/(1 - hc_1 f_x)$$

yields

$$f(x,\mu) = \frac{2+\mu}{h(2c_1 + c_1\mu - 1)}x + \frac{\tilde{s}}{h(2c_1 + c_1\mu - 1)(6c_1 + 3c_1\mu - 1)}x^3 + O(x^4).$$

Define

$$\alpha = \alpha(\mu) = \frac{\mu}{h(2c_1 + c_1\mu - 1)(2c_1 - 1)}.$$

Then $\alpha(0) = 0$ and $\alpha'(0) > 0$ for $c_1 \neq 1/2$, so a unique smooth local inverse function $\mu(\alpha)$ exists by the Inverse Function Theorem, and the function and its inverse monotonically increase in a neighbourhood of the origin. We have

$$f = (\alpha_0 - \alpha)x + \bar{s}a_3(\alpha)x^3 + O(x^4), \qquad \alpha_0 = \frac{2}{h\,(2c_1 - 1)},$$

where $a_3(\alpha)$ is smooth with $a_3(0) = [h(2c_1 - 1)(6c_1 - 1)]^{-1} \neq 0$, but $a_3(0)$ is singular for $c_1 = 1/6$ and $c_1 = 1/2$.

Finally, let $\eta = \sqrt{|a_3(\alpha)|}\,x$, yielding

$$\eta' = (\alpha_0 - \alpha)\eta + s\eta^3 + O(\eta^4),$$

where $s = \mathrm{sgn}(\bar{s}a_3(0)) = \mathrm{sgn}(\bar{s}(2c_1 - 1)(6c_1 - 1))$. Higher degree terms can eliminated due to local topological equivalences [2]. The origin $(x, \mu) = (0, 0)$ is preserved under the transformations $(x, \mu) \to (\eta, \alpha)$. Therefore, we interpret this as a normal form for scalar flow causing a spurious period doubling bifurcation at the origin under linearized one-point collocation methods. □

The normal forms of scalar vector fields causing spurious period doubling under collocation methods are $f(x, \alpha) = (\alpha_0 - \alpha)\,x \pm x^3$, where $\alpha_0 = \dfrac{2}{h\,(2c_1 - 1)}$. We can explicitly calculate the family of maps arising from the transformation by considering

$$g = x + \frac{h\,(\alpha_0 x - \alpha x + \bar{s}x^3)}{1 - hc_1(\alpha_0 - \alpha + 3\bar{s}x^2)}.$$

Fixed points are $x^* = 0$ and $x^* = \pm\sqrt{\bar{s}(\alpha - \alpha_0)}$. $x^* = 0$ is stable for:

$$\begin{cases} \dfrac{-2}{h\,(1-2c_1)} \overset{P}{<} \alpha \overset{PD}{<} 0 & (c_1 < 1/2) \\ \alpha \overset{PD}{<} 0 & (c_1 > 1/2) \\ \alpha \overset{P}{>} \dfrac{2}{h\,(2c_1 - 1)} & (c_1 > 1/2). \end{cases}$$

$x^* = \pm\sqrt{\bar{s}(\alpha - \alpha_0)}$ are stable for:

$$\begin{cases} \dfrac{-3}{h\,(1-2c_1)} < \alpha \overset{P}{<} \dfrac{-2}{h\,(1-2c_1)} & (c_1 < 1/2, \bar{s} = -1) \\ \alpha \overset{P}{<} \dfrac{2}{h\,(2c_1 - 1)} & (c_1 > 1/2, \bar{s} = -1) \\ \alpha > \dfrac{3}{h\,(2c_1 - 1)} & (c_1 > 1/2, \bar{s} = 1). \end{cases}$$

In the subcritical case, the nontrivial fixed points are $x^* = \pm\sqrt{-\alpha}$, always unstable when $c_1 < 1/2$, and stable for:

$$\left\{ \alpha \overset{PD}{<} \dfrac{-1}{h\,(2c_1 - 1)} \qquad (c_1 > 1/2). \right.$$

The labelled inequalities represent the locations of pitchfork (P) and period doubling (PD) bifurcations. See Figures 3.7 to 3.10.

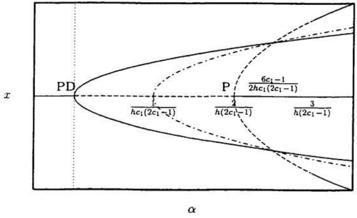

Figure 3.7: Period doubling: $f = (\alpha_0 - \alpha)x + x^3$, $\alpha_0 = \frac{2}{h(2c_1-1)}$, $c_1 > 0.5$

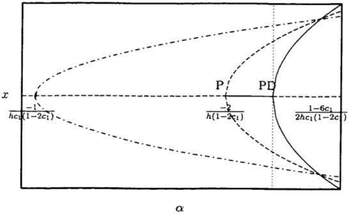

Figure 3.8: Period doubling: $f = (\alpha_0 - \alpha)x + x^3$, $\alpha_0 = \frac{2}{h(2c_1-1)}$, $0 < c_1 < \frac{1}{6}$

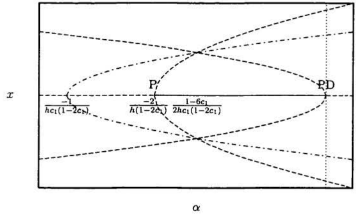

Figure 3.9: Period doubling: $f = (\alpha_0 - \alpha)x + x^3$, $\alpha_0 = \frac{2}{h(2c_1-1)}$, $\frac{1}{6} < c_1 < \frac{1}{2}$

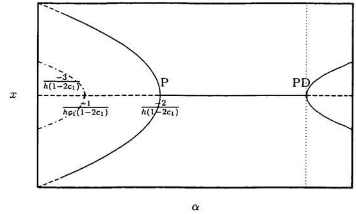

Figure 3.10: Period doubling: $f = (\alpha_0 - \alpha)x - x^3$, $\alpha_0 = \frac{2}{h(2c_1-1)}$, $\frac{1}{6} < c_1 < \frac{1}{2}$

The nontrivial fixed points exist for $\alpha > 0$ in the supercritical case, and for $\alpha < 0$, $\alpha \neq -1/(2hc_1)$, in the subcritical bifurcation. The trivial fixed point exists for all $\alpha \neq 1/(hc_1)$. The limiting case $c_1 = 0$ merely removes the singularity set to $\alpha \to +\infty$. The limiting case $c_1 = 1/2$ removes the period doubling bifurcations to $\alpha \to \pm\infty$.

3.2.5 Conclusion

It has been established that linearized one-point collocation methods do not admit spurious codimension-1 bifurcations, with the exception of the period doubling bifurcation which can occur in the methods for $c_1 \neq 1/2$. A normal form for the flow that gives rise to spurious period doubling bifurcations under transformation by these methods has been obtained.

It is worth noting, though, that while local dynamical behaviour of the methods may resemble that of the differential equation in the neighbourhood of the bifurcation, the presence of the singular set may distort the basins of attraction for the fixed points. This is discussed in Chapter 4.

3.3 Hopf and Neimark-Sacker Bifurcations

In this section, we assume $x' = f(x, \mu)$ for some $\mu \in \mathbb{R}$ is a C^k ($k \geq 3$) planar vector field. Isolating the linear part of f, we obtain

$$x' = J(\mu)x + F(x, \mu).$$

Assume the linear part $J(\mu)$ at x^* (x^* may depend on μ) has eigenvalues $\lambda_{1,2}(\mu) = \alpha(\mu) \pm i\beta(\mu)$. Assume further that $\alpha(\mu^*) = 0$ and $\beta(\mu^*) \neq 0$ and that for sufficiently small $|\mu - \mu^*|$, $F(x^*, \mu) = 0$ and $F_x(x^*, \mu) = 0$. Furthermore, suppose $\alpha'(\mu^*) \neq 0$. Then, the vector field undergoes a *Hopf bifurcation*, and in any neighbourhood U of x^* in $\mathbb{R}^2$ there exists with any given $\mu_0 > 0$, a $\tilde{\mu}$ with $|\tilde{\mu} - \mu^*| < \mu_0$ such that $x' = J(\tilde{\mu})x + F(x, \tilde{\mu})$ has a nontrivial periodic orbit in U.

3.3.1 Bifurcation Values in Differential Equations

In (3.1), let $x = (x_1, x_2)^T$ and $f(x, \mu) = (f_1(x_1, x_2, \mu), f_2(x_1, x_2, \mu))^T$. We define the quantities

$$A(\mu, x) = \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} \tag{3.12}$$

$$B(\mu, x) = \frac{\partial f_1}{\partial x_1}\frac{\partial f_2}{\partial x_2} - \frac{\partial f_1}{\partial x_2}\frac{\partial f_2}{\partial x_1}. \tag{3.13}$$

Then, the eigenvalues of $f_x(x^*(\mu), \mu)$ are

$$\begin{aligned} \lambda_1(\mu) &= \frac{1}{2}(a + \sqrt{a^2 - 4b}) \\ \lambda_2(\mu) &= \frac{1}{2}(a - \sqrt{a^2 - 4b}), \end{aligned}$$

where $a = A(x^*(\mu), \mu)$ and $b = B(x^*(\mu), \mu)$. If we let $\alpha = a/2$, then

$$\lambda_1 = \alpha + \sqrt{\alpha^2 - b} \tag{3.14}$$

$$\lambda_2 = \alpha - \sqrt{\alpha^2 - b}. \tag{3.15}$$

A necessary condition for the equilibrium point $x^*(\mu)$ to undergo a Hopf bifurcation at $\mu = \mu^*$ is $\alpha(\mu^*) = 0$ and $b(\mu) > \alpha^2$ in a neighbouhood of μ^*.

3.3.2 Bifurcation Values in the Numerical Methods

Now, let us assume that a numerical scheme such as (3.2) is used to solve (3.1) numerically. Assume g is a C^4 map and satisfies the conditions

(i) $g(\mu, x^*) = 0$ for μ near some fixed $\overline{\mu}$ (x^* may depend on μ);

(ii) $g_x(\mu, x^*)$ has two non-real eigenvalues $\sigma_{1,2}(\mu)$ for μ near $\overline{\mu}$ with $|\sigma(\overline{\mu})| = 1$;

(iii) $\dfrac{d}{d\mu}|\sigma(\mu)| > 0$ at $\mu = \overline{\mu}$;

(iv) $\sigma^k(\overline{\mu}) \neq 1$, for $k = 1, 2, 3, 4$.

Then, there is a neighbourhood U of x^* and a $\delta > 0$ such that, for $|\mu - \overline{\mu}| < \delta$ and $x_0 \in U$, either

(a) the omega limit set of x_0 is x^* if $\mu < \overline{\mu}$ and for $\mu > \overline{\mu}$, the omega limit set of x_0 is $\Gamma(\mu)$ where $\Gamma(\mu)$ is a closed invariant C^1 curve encircling x^*. Furthermore, $\Gamma(\overline{\mu}) = x^*$;

or,

(b) the alpha limit set of x_0 is x^* if $\mu > \overline{\mu}$ and for $\mu < \overline{\mu}$, the alpha limit set of x_0 is $\Gamma(\mu)$ where $\Gamma(\mu)$ is a closed invariant C^1 curve encircling x^*. Furthermore, $\Gamma(\overline{\mu}) = x^*$.

The map undergoes a *Neimark-Sacker* bifurcation, which is the discrete analog of the Hopf bifurcation in vector fields. In (a) above, the bifurcation is said to be *supercritical* and in (b) it is said to be *subcritical.*

We then have the following lemma.

Lemma 3.3.1 *For $\mu = \overline{\mu}$, let $\lambda_1(\overline{\mu})$ and $\lambda_2(\overline{\mu})$ be eigenvalues of $f_x(\overline{\mu}, x^*(\overline{\mu}))$. Then, the eigenvalues corresponding to the linearization of the methods (3.2) at x^* for $\mu = \overline{\mu}$ are given by*

$$\sigma_1(\overline{\mu}) = 1 + \frac{h\lambda_1(\overline{\mu})}{1 - hc_1\lambda_1(\overline{\mu})}, \tag{3.16}$$

$$\sigma_2(\overline{\mu}) = 1 + \frac{h\lambda_2(\overline{\mu})}{1 - hc_1\lambda_2(\overline{\mu})}. \tag{3.17}$$

Proof: Perform standard perturbation analysis on (3.2) about the fixed point $x^*(\overline{\mu})$, that is, define vectors δ^n and δ^{n+1} such that $X^{n+1} = x^*(\overline{\mu}) + \delta^{n+1}$ and $X^n = x^*(\overline{\mu}) + \delta^n$. Linearizing about $x^*(\overline{\mu})$, and dropping second and higher order terms in δ and h, we obtain

$$\delta^{n+1} = \delta^n A(x^*, \overline{\mu}), \tag{3.18}$$

where $A(x^*, \overline{\mu}) = I + h(I - hc_1 f_x(x^*, \overline{\mu}))^{-1} f_x(x^*, \overline{\mu})$.

If the eigenvalues of $f_x(x^*, \overline{\mu})$ are $\lambda_1(\overline{\mu})$ and $\lambda_2(\overline{\mu})$, then the eigenvalues of $A(x^*, \overline{\mu})$ are given by (3.16) and (3.17). □

We now consider the problem of determining whether solving a system with a Hopf bifurcation by linearized one-point collocation methods will result in the same bifurcation occurring at the same parameter value. The following theorem answers this question.

Theorem 3.3.2 *Let (3.1) be a system that undergoes a Hopf bifurcation at the parameter value $\mu = \mu^*$. We represent the eigenvalues of the Jacobian matrix of f at any equilibrium point $(\mu, x^*(\mu))$ where μ is sufficiently close to μ^*, by $\lambda_{1,2} = \alpha \pm \sqrt{\alpha^2 - b}$ where $b > \alpha^2$. Then, if a linearized one-point collocation method with fixed stepsize is used to discretize the system, the method will exhibit a Neimark-Sacker bifurcation, occurring at a parameter value $\mu = \overline{\mu}$, where*

$$2\alpha(\overline{\mu}) = hb(\overline{\mu})(2c_1 - 1). \tag{3.19}$$

Proof: Assume the non-degeneracy conditions are satisfied. Making use of (3.16) and (3.17), and performing some algebraic manipulations, we obtain

$$\sigma_{1,2} = \left(1 + \frac{bh(\alpha - bhc_1)}{b - 2bhc_1\alpha + b^2h^2c_1^2}\right) \pm \frac{bh}{b - 2bhc_1\alpha + b^2h^2c_1^2}\sqrt{\alpha^2 - b}. \tag{3.20}$$

Now, $|\sigma_{1,2}(\overline{\mu})| = 1$ if either

$$b(\overline{\mu})[1 - 2hc_1\alpha(\overline{\mu}) + b(\overline{\mu})h^2c_1^2] = 0$$

or

$$2\alpha(\overline{\mu}) - b(\overline{\mu})h\{2c_1 - 1\} = 0.$$

The first of the above conditions represents a singular point for the eigenvalues, and the second yields

$$2\alpha(\overline{\mu}) = hb(\overline{\mu})[2c_1 - 1]. \tag{3.21}$$

This completes the proof. □

Under our assumptions, it is clear that $\overline{\mu} = \mu^*$ if and only if $c_1 = \frac{1}{2}$. Hence, the following corollary.

Corollary 3.3.3 *The Neimark-Sacker bifurcation occurs for the same parameter value as with the Hopf bifurcation in the differential equation if and only if $c_1 = 1/2$, that is, for the linearized implicit midpoint method.*

3.4 Examples

3.4.1 Van der Pol's Oscillator (Unforced)

The Van der Pol oscillator is given by

$$x'' + \gamma(x^2 - 1)x' + x = 0. \tag{3.22}$$

where γ is a parameter. This system can be transformed to the 2-D system

$$\begin{cases} x_1' &= x_2 \\ x_2' &= -x_1 - \gamma(x_1^2 - 1)x_2. \end{cases} \tag{3.23}$$

Clearly, the origin is the only equilibrium point of the system (3.23). The eigenvalues of the linearization of the system about the equilibrium point are given by

$$\lambda_{1,2} = \frac{\gamma \pm \sqrt{\gamma^2 - 4}}{2}. \tag{3.24}$$

If we make the transformation $\epsilon = \gamma/2$, then

$$\lambda_{1,2} = \epsilon \pm \sqrt{\epsilon^2 - 1}. \tag{3.25}$$

We assume $-1 < \epsilon < 1$, so that the eigenvalues are complex. The origin undergoes a supercritical Hopf bifurcation at $\epsilon = 0$.

For one-point collocation methods we have, according to Theorem 3.3.2, the Neimark-Sacker bifurcation at the parameter value

$$\overline{\epsilon} = \frac{h}{2}(2c_1 - 1).$$

3.4.2 Forced Oscillators and the Method of Averaging

The *method of averaging* transforms a nonautonomous differential equation to an autonomous one. Suppose we have a differential equation of the form

$$x' = \epsilon f(x,t) + \epsilon^2 g(x,t,\epsilon) \tag{3.26}$$

where $f : U \times \mathbb{R} \to \mathbb{R}^m$ and $g : U \times \mathbb{R} \times [0, \epsilon_0) \to \mathbb{R}^m$ are C^r $(r \geq 1)$ on their respective domains of definition and T-periodic $(T > 0)$ in t. Here, U is an open set in $\mathbb{R}^m$. Then, the averaged equation will take the form

$$y' = \epsilon \overline{f}(y), \qquad y \in \mathbb{R}^m, \tag{3.27}$$

where

$$\overline{f}(y) = \frac{1}{T}\int_0^T f(y,t)dt.$$

Starting with a differential equation of the form

$$x' = f(x) + \epsilon g(x,t,\epsilon), \qquad x \in \mathbb{R}^m, \tag{3.28}$$

we can transform it to the form (3.26) by performing the simple steps (see Wiggins, 1990 for details):

- Consider a solution, $y(t) \equiv x(t, x_0(t))$, of (3.28), taking the initial condition as a function of t. Differentiating with respect to t and rearranging gives $x_0' = (D_{x_0}x)^{-1}[f(x(t,x_0)) - x' + \epsilon g(x(t,x_0),t,\epsilon)]$.

- If $x(t, x_0)$ is a solution of (3.28) with $\epsilon = 0$ (unperturbed equation), the above reduces to

$$x_0' = \epsilon (D_{x_0} x)^{-1} g(x(t, x_0), t, \epsilon)).$$

The dynamics of the averaged equation is related to that of the original equation by the following theorem from Guckenheimer & Holmes [11].

Theorem 3.4.1 *Consider the differential equation*

$$x' = \epsilon f(x, t; \mu) + \epsilon^2 g(x, t, \epsilon; \mu) \tag{3.29}$$

and its associated averaged equation

$$y' = \epsilon \overline{f}(y, \mu), \qquad \overline{f} = \frac{1}{T} \int_0^T f(y, t; \mu) dt, \tag{3.30}$$

where $\mu \in \mathbb{R}$ is a parameter.

If, at $\mu = \mu^*$, (3.30) undergoes a saddle-node or Hopf bifurcation, then, for μ near μ^* and ϵ sufficiently small, the Poincaré map of (3.29) undergoes a saddle-node or Neimark-Sacker bifurcation.

Van der Pol's Oscillator (Forced)

The forced Van der Pol's oscillator

$$x'' + \frac{\epsilon}{\omega}(x^2 - 1)x' + x = \epsilon F \cos \omega t \tag{3.31}$$

is transformed by the method of averaging to the autonomous system

$$\begin{aligned} u' &= u - \sigma v - u(u^2 + v^2) \\ v' &= \sigma u + v - v(u^2 + v^2) - \gamma, \end{aligned} \tag{3.32}$$

where $\epsilon\sigma = 1 - \omega^2$ and $\gamma = F/2$ (See Guckenheimer & Holmes [11]).

Figure 3.11 depicts the bifurcation diagram for (3.32). A Hopf bifurcation occurs along the curve marked OE, whose equation is given by $8\gamma^2 = 4\sigma^2 + 1$ where $|\sigma| > 0.5$.

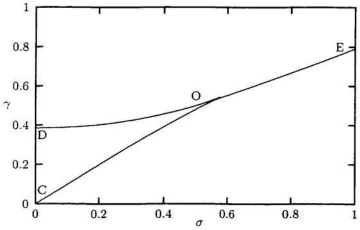

Figure 3.11: Bifurcation diagram of the averaged equations (3.32)

Guckenheimer & Holmes [11] discuss the diagram in more detail from the point of view of continuous dynamics. Our interest here will be to study the effect of discretization by using linearized one-point collocation methods on the location of the Hopf bifurcation. We take a cross-section of the diagram at $\sigma = 0.75$, and vary γ. The Hopf bifurcation (subcritical) then occurs at $\gamma^* = \sqrt{13/32} \approx 0.6373774$. It can be shown that for any value of γ and σ, the equilibrium points of the system (3.32) satisfy

$$\gamma^2 u^3 - 2\sigma\gamma u^2 + \sigma^2(\sigma^2+1)u - \gamma\sigma^3 = 0 \quad (3.33)$$

$$v = \tfrac{1}{\sigma}\left(u - \tfrac{\gamma}{\sigma}u^2\right). \quad (3.34)$$

Recall that the eigenvalues of the linearization

$$J(u,v) = \begin{pmatrix} 1-3u^2-v^2 & -\sigma-2uv \\ \sigma-2uv & 1-3v^2-u^2 \end{pmatrix}$$

of the function on the right of (3.32) at an equilibrium point determined by γ ($\sigma = 0.75$) are given by

$$\lambda_{1,2}(\gamma) = \alpha(\gamma) \pm \sqrt{b-\alpha^2}i; \qquad b > \alpha^2,$$

where $\alpha(\gamma^*) = 0$. The canonical form of the Jacobian is then given by

$$J(u^*(\gamma), v^*(\gamma)) = \begin{pmatrix} \alpha(\gamma) & \alpha^2(\gamma) - b(\gamma) \\ 1 & \alpha(\gamma) \end{pmatrix}.$$

Recall also that the Neimark-Sacker bifurcation in the collocation methods will occur at $\gamma = \overline{\gamma}$, where

$$2a(\overline{\gamma}) = b(\overline{\gamma})h(2c_1 - 1). \tag{3.35}$$

On the other hand, we observe that

$$\text{Det}[J(u^*(\overline{\gamma}, v^*(\overline{\gamma})] = (1 - 3u^2 - v^2)(1 - 3v^2 - u^2) + \sigma^2 - 4u^2v^2 = b(\overline{\gamma}) \tag{3.36}$$

and

$$\text{Trace}[J(u^*(\overline{\gamma}, v^*(\overline{\gamma})] = 2a(\overline{\gamma}) = 2 - 4u^2 - 4v^2. \tag{3.37}$$

We solve the equations (3.33), (3.34), (3.36) and (3.37) simultaneously using (3.35) on MAPLE to obtain (u^*, v^*) and the corresponding $\overline{\gamma}$. In Figure 3.12, we plot the bifurcation values $\overline{\gamma}$ for various values of c_1 and $h = 0.01$. Note that, for $c_1 < 0.5$, the numerical bifurcation values are below the true value and the opposite is true for $c_1 > 0.5$.

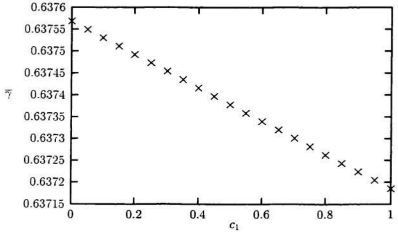

Figure 3.12: Bifurcation values of collocation methods for the averaged equations of the forced VDP oscillator.

Chapter 4

Spurious Poles in Numerical Solutions to Autonomous Equations

4.1 Introduction

In this chapter, we consider the possible effect of the singularity in linearized one-point collocation methods on the dynamics of autonomous equations. Recall that, applied to the autonomous ODE (1.2), the linearized one-point collocation methods are given by

$$x_{n+1} = x_n + h[I - hc_1 f_x(x_n)]^{-1} f(x_n) \tag{4.1}$$

where f_x is the Jacobian of f. The existence of the inverse of the matrix $I - hc_1 f_x$ in the method may result in singularities. We will first demonstrate what effect this could have on the asymptotic behaviour of the solution, then discuss some general theory on singularity analysis and poles.

The discussion begins with a consideration of global dynamics in some of the normal form equations in § 3.2 of Chapter 3.

4.2 Normal Form Equations: Codimension-1 Bifurcations

4.2.1 Saddle-node Bifurcation

Recall that a normal form for the saddle-node bifurcation is given by

$$x' = \alpha + x^2.$$

Hence the map

$$x \mapsto x + \frac{h(\alpha + x^2)}{1 - 2hc_1 x}$$

represents an application of linearized one-point collocation methods to the normal form.

The case $c_1 = 1/2$ is of special interest since it produces a locally superconvergent method. Taking $c_1 > 1/2$, we use linear stability analysis to establish that a local period-2 cycle exists for $\alpha > \frac{-1}{h^2(2c_1 - 1)^2}$. This cycle is stable for

$$\frac{-1}{h^2(2c_1 - 1)^2} < \alpha < \frac{3 - 10c_1}{2h^2(2c_1 - 1)(10c_1^2 - 6c_1 + 1)}.$$

For these values of α, the cycle attracts all x-values such that $\frac{1}{2hc_1} < x < \infty$.

The system goes through a period doubling cascade to chaos, before a boundary crisis bifurcation occurs at $\alpha = \dfrac{-1}{2h^2c_1(2c_1 - 1)}$. A boundary crisis is created when the basin boundary of a chaotic attractor collides with an unstable fixed point. In this case, the unstable "fixed point" is the singular set. As a result of the boundary crisis, the chaotic set is no longer attracting and becomes transient. For $\dfrac{-1}{2h^2c_1(2c_1 - 1)} < \alpha < 0$, all initial values, apart from the singular set, are mapped to the stable fixed point $x^* = -\sqrt{-\alpha}$. In this range, the numerical method turns the stable fixed point into a global attractor, which is not the case in the ODE. In the ODE, the stable equilibrium attracts all $x < \sqrt{-\alpha}$, and for all $x > \sqrt{-\alpha}$, orbits go to ∞. Furthermore, the monotonicity of the orbits for $x > \sqrt{-\alpha}$ is lost in the numerical methods. This is because the direction of orbits destined for $+\infty$ is changed when they cross the singular set. For $\alpha > 0$, the system has no fixed points and the chaotic set, which has been transient, reappears as a type I intermittency (see Pomeau & Manneville [25], Foster [8]). Figure 4.1 is the bifurcation diagram for the system with $h = 0.1$ and $c_1 = 1$.

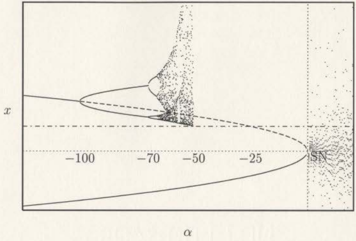

Figure 4.1: Saddle-node: $f = \alpha + x^2$, $c_1 = 1$, $h = 0.1$

4.2.2 Transcritical Bifurcation

A normal form for the transcritical bifurcation is given by

$$x' = \alpha x + x^2.$$

Discretizing using linearized one-point collocation methods results in the map

$$x \mapsto x + \frac{h(\alpha x + x^2)}{1 - hc_1(\alpha + 2x)}.$$

Using $c_1 > 1/2$, we establish that there is a local period-2 cycle for $\alpha^2 < \frac{4}{h^2(2c_1-1)^2}$. This cycle is stable (and attracts all initial values above the singular set) for

$$\frac{-2}{h(2c_1 - 1)} < \alpha < -\sqrt{\frac{2(10c_1 - 3)}{h^2(20c_1^3 - 22c_1^2 + 8c_1 - 1)}}$$

and

$$\sqrt{\frac{2(10c_1 - 3)}{h^2(2c_1 - 1)(10c_1^2 - 6c_1 + 1)}} < \alpha < \frac{2}{h(2c_1 - 1)}.$$

The system undergoes period doubling cascades to chaos, culminating with boundary crisis bifurcations at

$$\alpha = \pm\sqrt{\frac{2}{h^2 c_1 (2c_1 - 1)}},$$

and for $|\alpha| < \sqrt{\frac{2}{h^2 c_1 (2c_1 - 1)}}$, all initial conditions apart from the singular set itself yield orbits that converge to the stable fixed point. Again, the map converts the fixed point into a global attractor, and the singular set destroys the monotonicity of orbits. Figure 4.2 is the complete bifurcation diagram for this map.

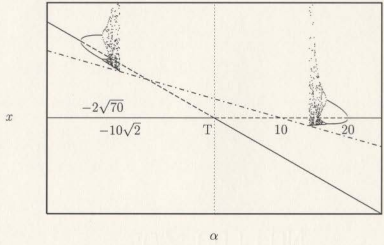

Figure 4.2: Transcritical: $f = \alpha x + x^2$, $c_1 = 1$ and $h = 0.1$.

4.2.3 Subcritical Pitchfork Bifurcation

A normal form for the subcritical pitchfork bifurcation is given by

$$x' = \alpha x + x^3,$$

hence the map

$$x \mapsto x + \frac{h(\alpha x + x^3)}{1 - hc_1(\alpha + 3x^2)} \tag{4.2}$$

represents an application of linearized one-point collocation methods to the normal form.

Due to the complexity of the equations for the locations of the bifurcations in the map (4.2), we let $h = 0.1$ and $c_1 = 1$.

Figure 4.3 shows the complete bifurcation diagram for (4.2). As we saw with the saddle-node and transcritical normal forms, there are spurious period-doubling bifurcations, period-doubling cascades and chaotic behaviour. In addition, at $\alpha = 90/7$, a period-2 cycle undergoes a transcritical bifurcation, then a period-doubling bifurcation at $\alpha = \dfrac{20+50\sqrt{5}}{11}$. We identify Type III intermittent behaviour (Pomeau & Manneville [25]) to the right of the pitchfork bifurcation.

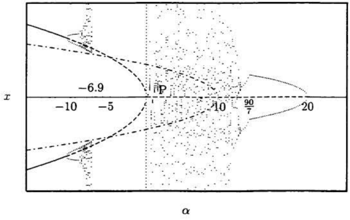

Figure 4.3: Subcritical Pitchfork: $f = \alpha x + x^3$, $c_1 = 1$, $h = 0.1$

4.2.4 Remarks

From the preceding examples, it is clear that the global dynamics of the methods is different from that of the corresponding system. Not only do the methods give rise to spurious period doubling and chaotic behaviour, but they change the asymptotics of trajectories and drastically alter the basins of attraction of the fixed points. However, local to a bifurcation, the local dynamics and asymptotics of the ODE are preserved. Typically, in the presence of a stable fixed point and an unstable one, the ω-limit set of any orbit with initial condition not on the singular set is the stable fixed point; the exception being the values of the parameter for which the spurious period-doubling cascade is present. This does not occur in scalar flows, and is directly due to the presence of the singular set in the methods.

We further motivate the study of the effect of the singular set on global dynamics by considering a planar predator-prey system.

4.3 Motivation: A Predator-Prey System

Consider the two-dimensional system of nonlinear first-order autonomous ordinary differential equations,

$$\begin{cases} u' &= -3u + 4u^2 - 0.5uv - u^3 \\ v' &= -2.1v + uv \end{cases} \tag{4.3}$$

where u and v represent the population of the prey and predator respectively. The equilibrium points of this system are (0,0) (stable node), (2.1, 1.98) (stable spiral), (1, 0) and (3, 0) (saddles).

Yee & Sweby [33] studied the global asymptotic behaviour of the above system using a number of linear multistep methods. What they uncovered is that different numerical schemes can give rise to differing appearances of basins of attraction of local fixed points. In general, the numerical basins of attraction bore no resemblance to the exact basins of attraction. Depending on the discretization parameter, the numerical basins also differed significantly from each other (see Chapter 2). Our objective here is to explain this behaviour by appealing to the singular set analysis of the maps.

4.3.1 Linearized One-point Collocation Methods

Applying the methods (4.1) to the system (4.3) gives

$$
\begin{aligned}
u_{n+1} &= u_n + \frac{h(1+2.1hc_1 - hc_1u_n)(-3u_n + 4u_n^2 - 0.5u_nv_n - u_n^3)}{Q(u_n, v_n)} + \\
&\quad \frac{-0.5h^2c_1u_n(-2.1v_n + u_nv_n)}{Q(u_n, v_n)} \\
v_{n+1} &= v_n + \frac{h^2c_1v_n(-3u_n + 4u_n^2 - 0.5u_nv_n - u_n^3)}{Q(u_n, v_n)} + \\
&\quad \frac{h(1+3hc_1 - 8hc_1u_n + 0.5hc_1v_n + 3hc_1u_n^2)(-2.1v_n + u_nv_n)}{Q(u_n, v_n)}
\end{aligned}
\tag{4.4}
$$

where h is the time step and

$$Q(u_n, v_n) = (1+3hc_1-8hc_1u_n+0.5hc_1v_n+3hc_1u_n^2)(1+2.1hc_1-hc_1u_n)+0.5h^2c_1^2u_nv_n.$$

Performing some standard stability analysis of each of the four fixed points, we discover that, depending on c_1 and the stepsize h, the dynamics of the numerical methods may differ significantly from that of the system. The stability analysis was performed analytically and verified using MAPLE. Figures 4.4 to 4.7 are bifurcation diagrams for each of the four fixed points, as obtained analytically.

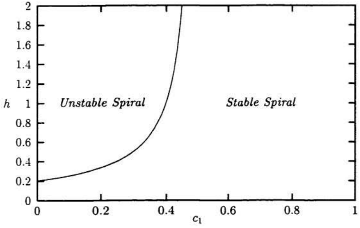

Figure 4.4: Bifurcation curve for the fixed point (2.1,1.98).

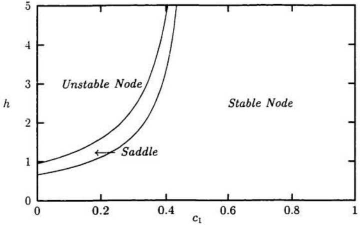

Figure 4.5: Bifurcation curve for the fixed point (0,0).

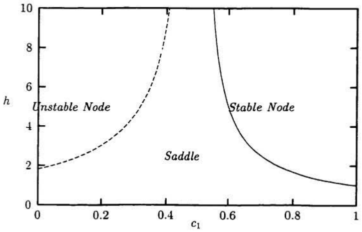

Figure 4.6: Bifurcation curve for the fixed point (1,0).

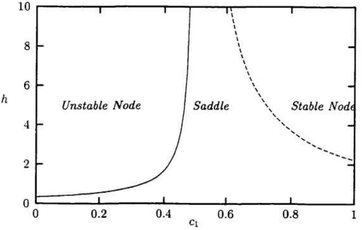

Figure 4.7: Bifurcation curve for the fixed point (3,0).

While the methods do not admit spurious fixed points, the diagrams show that they change the stability types. For example, the saddle $(1, 0)$ is converted to a stable node by a linearized one-point collocation method with $c_1 > 0.5$ and $h > \frac{1}{2c_1 - 1}$.

4.3.2 Basins of Attraction

Depicting basins of attraction (showing global asymptotic behaviour) is very revealing. Figure 4.8 shows the true basin of attraction for the equilibrium points of (4.3) as shown in Yee & Sweby [33]. Figure 4.9 shows the basin as computed by the linearized implicit midpoint method with $h = 0.1$. In both diagrams, red represents the basin for $(0, 0)$, blue for $(2.1, 1.98)$ and black for infinity. The darker shade of blue indicates a slower rate of convergence to $(2.1, 1.98)$.

The set of initial values that is attracted to $(0, 0)$ is larger in the numerical basin than in the true basin. This is due to the shrinking of the basin for ∞. The reason for this behaviour is the existence of the singular set $Q(u, v) = 0$. All orbits with initial value in the true basin for $-\infty$ are attracted to the fourth-quadrant subset of the unstable manifold, W^u, of the saddle at $(u, v) = (3, 0)$. However, because of the singular set, these orbits are "diverted" into the basin for $(0, 0)$ upon reaching some neighbourhood of the singular set. Such behaviour will be defined in the next section as *pole-type behaviour.*

For $h = 0.1$ and $c_1 = 0.5$, $E^u = \text{span}(-0.21739, 1)^T$. W^u is tangent to the vector E^u at $(3, 0)$.

Figure 4.10 shows the location of the singular set $Q(u, v) = 0$ of linearized one-point collocation methods with $hc_1 = 0.05$ for the system (4.3). The figure also shows one orbit of the discrete system with initial value, x_0, in the true basin of attraction of $-\infty$. For this initial condition, $\omega(x_0) = (0, 0)$. When the orbit, while moving along W^u, reaches a neighbourhood of the singular set, it loses its monotonicity, falls in the basin of $(0, 0)$ and then approaches this fixed point monotonically, as shown. The loss of monotonicity occurs at approximately $v = -800$ (not shown).

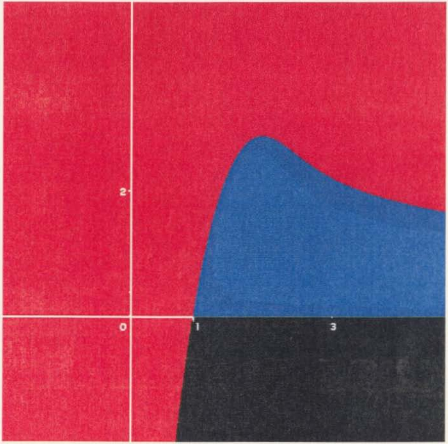

Fig 4.8: Basin of attraction - Predator-Prey Model

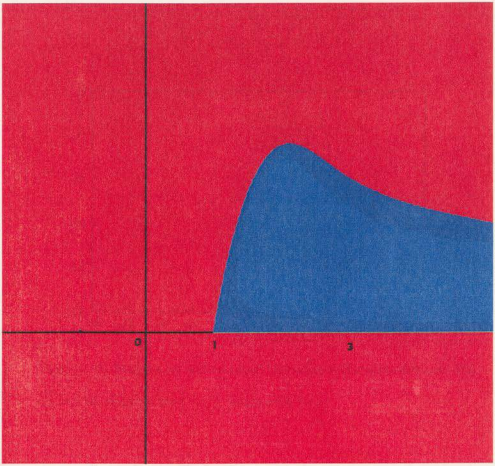

Fig 4.9: Numerical Basin of local fixed points (h=0.1, c_i=0.5)

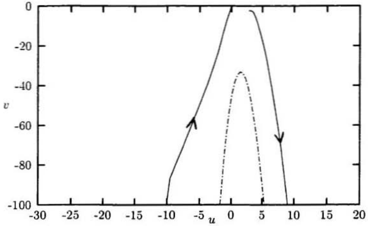

Figure 4.10: Singular set of Linearized One-point collocation methods with $hc_1 = 0.05$. Singular set (- · · ·), single orbit (—).

4.4 Spurious Pole-type Behaviour

Linearized one-point collocation methods, formulated specifically for the planar autonomous equation $u' = f_1(u,v)$, $v' = f_2(u,v)$ can be written in the form

$$u_{n+1} = u_n + h\frac{P(u_n, v_n)}{Q(u_n, v_n)}$$
$$v_{n+1} = v_n + h\frac{R(u_n, v_n)}{Q(u_n, v_n)} \tag{4.5}$$

where $P(u, v)$, $Q(u, v)$ and $R(u, v)$ are functions involving the f_1 and f_2 and their partial derivatives.

Definition 4.4.1 *We define the singular set of the method as the subset of* $\mathbb{R}^2$ *where* $Q(u, v) = 0$.

All linearized one-point collocation methods with $c_1 > 0$ have a nontrivial singular set. The location and nature of the singular set is important since we expect the dynamics of the method near the singular set to be unpredictable, so that orbits destined for infinity are "disrupted" and end up finite. This behaviour is analogous to the existence of poles in continuous systems.

Definition 4.4.2 (Albowitz & Segur [1]) *Consider any ordinary or partial differential equation. Singularities of the coefficients of the differential equation will also be singularities of the solution. Such singularities are called* fixed singularities. *Any singularity of the solution that is initial condition dependent is called a* movable singularity.

Definition 4.4.3 *A differential equation is said to possess the Painlevé property (PP) if all movable singularities of its solutions are poles.*

Joshi [21] alluded to the existence of behaviour analogous to poles in discrete dynamical systems. The following is an adaptation of his definition to $\mathbb{R}^m$.

Definition 4.4.4 (Joshi [21]) *An orbit of a discrete dynamical system that begins in the finite part of* $\mathbb{R}^m$*, reaches the point at infinity in a finite number of steps* n*, and continues past infinity to the finite region of* $\mathbb{R}^m$ *is said to have* pole-type *behaviour.*

The integration of a differential equation that has no PP by a numerical method that exhibits pole-type behaviour can result in distortion or misrepresentation of the global asymptotic behaviour (basin of attraction).

Consider a scalar differential equation of the form

$$x' = P(x), \tag{4.6}$$

where $P(x)$ is a polynomial of degree $q > 2$ in x. This differential equation has no PP ([21]).

Theorem 4.4.5 *A linearized one-point collocation method with* $c_1 > 0$*, applied to (4.6), exhibits spurious pole-type behaviour if and only if* $c_1 = 1/q$.

Proof: If the numerical methods are represented by the map

$$x_{n+1} = x_n + h\phi(x_n, h, c_1),$$

then it suffices to show that $\lim_{x_n \to \infty} x_{n+1} = d < \infty$.

Let

$$P(x) = \sum_{j=0}^{q} a_j x^j.$$

Then, applying one-point collocation methods to (4.6), gathering terms and simplifying, gives

$$\begin{aligned} x_{n+1} &= \frac{x_n + h\sum_{j=0}^{q}(1 - jc_1)a_j x_n^j}{1 - hc_1\sum_{j=1}^{q} j a_j x_n^{j-1}} \\ &= \frac{x_n^{1-q} + h\sum_{j=0}^{q}(1 - jc_1)a_j x_n^{j-q}}{x_n^{-q} - hc_1\sum_{j=1}^{q} j a_j x_n^{j-q-1}}. \end{aligned}$$

Clearly, if $c_1 \neq 1/q$, $\lim_{x_n \to \infty} x_{n+1} = \infty$.

If $c_1 = 1/q$,

$$\begin{aligned} x_{n+1} &= \frac{qx_n + h\sum_{j=0}^{q-1}(q - j)a_j x_n^j}{q - h\sum_{j=1}^{q} j a_j x_n^{j-1}} \\ &= \frac{qx_n^{2-q} + h\sum_{j=0}^{q-1}(q - j)a_j x_n^{j-q+1}}{qx_n^{-q+1} - h\sum_{j=1}^{q} j a_j x_n^{j-q}}, \end{aligned}$$

so that

$$\lim_{x_n \to \infty} = -\frac{a_{q-1}}{qa_q} < \infty,$$

and the conclusion of the theorem follows. □

4.4.1 Remarks

In spite of the fact that we have not proved the existence of spurious pole-type behaviour in systems in which $f(x)$ is not a polynomial, we remark here that Theorem 4.4.5 and the results of the first two sections of this chapter, suggest that the phenomenon of spurious pole-type behaviour does occur in the general case. Indeed, if f is approximated by a truncated power series, spurious pole-type behaviour will be observed for some values of c_1.

Chapter 5

Nonautonomous Equations and Periodic Solutions

5.1 Introduction

Let

$$x' = f(x, t), \qquad x(0) = x_0 \tag{5.1}$$

where $f : I \subset \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, be a scalar ordinary differential equation in which $f(x, t)$ is a periodic function of t with prime period T.

Since f depends explicitly on t, the differential equation is said to be *nonautonomous*. The detailed dynamics of numerics for nonautonomous ODEs has notably been lacking until the work in Khumalo [22] which is based on the material in this chapter.

Although any nonautonomous ODE can be transformed to an autonomous one, thereby increasing the dimension by one, the familiar dynamics of autonomous equations which is centered around the notion of equilibrium points, is lost. In certain special cases, this notion is replaced by that of periodicity. It is on these special cases that we will focus our attention. Stuart [30], proved using bifurcation theory that for reaction-diffusion-convection equations, linearized instability implies the existence of spurious periodic solutions. Our consideration here concentrates on nonautonomous ODEs where f is periodic in t, and differs from that of Stuart who considered partial differential equations. Nonautonomous ODEs with periodic solutions are very common in applications such as population dynamics with seasonal parameters or periodically forced systems.

Under certain conditions on f, (1.17) has a unique T-periodic solution. These will be stated for each form of f we will consider. We will assume the solution is approximated by a linearized one-point collocation method. Our objective is to determine, for each f under consideration, whether the numerical scheme has the same dynamical behaviour as the differential equation. In particular, we will consider cases in which the ODE has a unique, asymptotically stable periodic solution and establish conditions under which the numerical methods have the same dynamics. These special cases will take the following form:

(i) f linear: $f(x,t) = a(t)x + b(t)$, where $a(t)$ and $b(t)$ are C^1 T-periodic functions of t.

(ii) f nonlinear: $f = a(t)g(x) + b(t)$, where $a(t)$ and $b(t)$ are C^1 T-periodic functions of t and $g(x)$ is a C^2 nonlinear function of x.

Recall that the linearized one-point collocation methods for (1.17) are given by

$$x_{n+1} = \frac{x_n + hf(t_n, x_n) + c_1 h^2 \frac{\partial f}{\partial t}(t_n, x_n) - c_1 h x_n \frac{\partial f}{\partial x}(t_n, x_n)}{1 - c_1 h \frac{\partial f}{\partial x}(t_n, x_n)}. \tag{5.2}$$

We begin with a description of the dynamical systems theory approach, which will be used in determining the conditions under which the methods have the same dynamical behaviour as the differential equations. Upon establishing these conditions, we compare them with those imposed by nonautonomous stability theory.

5.1.1 Dynamical Systems Approach

In what follows, we will use a technique known as *stroboscopic sampling* to reduce the problem of determining existence and stability of periodic solutions to existence and stability of fixed points.

Let $x_{n+1} = p(x_n; n; h)$ be the discrete system representing the numerical method, applied with fixed stepsize, to the nonautonomous differential equation.

Step 1: Using inductive arguments, write the method in the form $x_{n+1} = \phi(x_0; n; h)$.

Step 2: Choose h such that the period, $T = hk$. Then, $x_k = \phi(x_0; k; h)$, and then we establish the new discrete system $X_{n+1} = \phi(X_n; k; h)$. This is known as stroboscopic sampling.

Step 3: The fixed points of the last system correspond to T-periodic solutions of the method. These are determined, with their stability types.

The above procedure is analogous to the Poincaré map of (1.17): Let $\Phi(t, x_0)$ be the T-periodic solution of (1.17), with starting value $x(0) = x_0$. Then, the Poincaré map of (1.17) is the scalar mapping

$$\Pi : \mathbb{R} \to \mathbb{R}; \qquad x \mapsto \Phi(T, x).$$

5.2 Linear Case

Suppose $f = a(t)x + b(t)$, where a and b are T-periodic functions of t. Then, the linear nonautonomous differential equation (1.17) becomes

$$x' = a(t)x + b(t), \qquad \text{for } t \geq 0 \text{ and } x(0) = x_0. \tag{5.3}$$

If $\nu_1 = \int_0^T a(t)dt < 0$, then (5.3) has a unique T-periodic asymptotically stable solution. If $a(t) = 0$, then (5.3) has a unique T-periodic solution that is asymptotically stable if $\nu_2 = \int_0^T b(s)ds = 0$ (Hale & Koçac, 1991).

Now, the linearized one-point collocation methods are given by

$$x_{n+1} = x_n + \frac{h}{1 - hc_1 a(nh)}\{a(nh)x_n + b(nh) + hc_1[a'(nh)x_n + b'(nh)]\}. \tag{5.4}$$

We will begin our discussion with the cases in which the function $a(t)$ is trivially periodic ($a(t) = 0$ and $a(t) = -1$).

5.2.1 Linear Case with $a(t) = 0$

Theorem 5.2.1 *Suppose a linearized one-point collocation method is used to solve the linear nonautonomous differential equation (5.3) with $a(t) = 0$. The method tends (as $n \to \infty$, $h > 0$ fixed) to a periodic solution for any starting value if and only if*

$$\sum_{r=0}^{k-1} \tilde{b}_r = 0,$$

where $\tilde{b}_r = b(rh) + hc_1 b'(rh)$ for each r.

Proof: Assume $\nu_2 = \int_0^T b(s)ds = 0$. Taking $a(t) = 0$ in (5.4), we obtain

$$x_{n+1} = x_n + h\{b(nh) + hc_1 b'(nh)\}.$$

We define

$$\tilde{b}_i = b(ih) + hc_1 b'(ih) \qquad \text{for } i = 0, 1, 2, \ldots \tag{5.5}$$

and denote

$$\Pi(x_n) = x_n + h\{b(nh) + hc_1 b'(nh)\}. \tag{5.6}$$

Now, we can prove by induction that for a given value of x_0,

$$x_{n+1} = x_0 + h \sum_{r=0}^{n} \bar{b}_r. \tag{5.7}$$

We can select h in such a way that T is an integral multiple of h. That is, we can fix $k \in \mathbf{N}$ such that $T = hk$. Consider the k-th iterate of x_0 under Π:

$$x_k = \Pi^k(x_0) = x_0 + \frac{T}{k} \sum_{r=0}^{k-1} \bar{b}_r \tag{5.8}$$

and the related iteration, which corresponds to *stroboscopic sampling*

$$X_{n+1} = X_n + \frac{T}{k} \sum_{r=0}^{k-1} \bar{b}_r \tag{5.9}$$

where $X_0 = x_0$. If the summation on the right-hand side of (5.9) is zero, then the discrete system is fixed at X_0 for all n, which corresponds to a periodic solution. If it is nonzero, then the stroboscopic iteration has no fixed point and diverges. □

Remark

The second term on the right-hand side of (5.9) can be viewed as an application of the left rectangular quadrature rule to the integral

$$\int_0^T (b(s) + c_1 \frac{T}{k} b'(s))\, ds.$$

From the assumption that $\nu_2 = 0$, we conclude that the above integral is zero, and (5.9) can be written as the simple map

$$X_{n+1} = X_n + c \tag{5.10}$$

where

$$|c| = \frac{T}{k} |\sum_{r=0}^{k-1} \bar{b}_r| = \frac{T^2}{2k} |b'(\xi) + c_1 \frac{T}{k} b''(\xi)| \tag{5.11}$$

for some $\xi \in (0, T)$, is the quadrature error.

The above theorem can then be restated as follows: Suppose a linearized one-point collocation method is used to solve the linear nonautonomous differential equation (5.3) with $a(t) = 0$. The method tends (as $n \to \infty$, $h > 0$ fixed) to a periodic solution for any starting value if and only if the rectangular quadrature rule, used to approximate the integral in (5.11), gives an exact result.

Suppose that $c \neq 0$ in the last theorem. We remark that the rate at which the system (5.9) grows or decays is dependent on the value of $|c|$, which, in turn depends on h. If h is very small (corresponding to large k), the growth/decay will be so small that in order for it to be significant, one would have to integrate over significantly long times.

Illustration

We examine the stroboscopic sampling of the solution of the differential equation $x' = \cos t e^{\sin t}$; $x(0) = 1$. Figure 5.1 shows the numerical results for $k = 4$, 5 and 10 with $c_1 = 0.5$. For $k = 4$, the method diverges quicker from the periodic solution than for $k = 5$. For $k = 10$, the divergence is negligible.

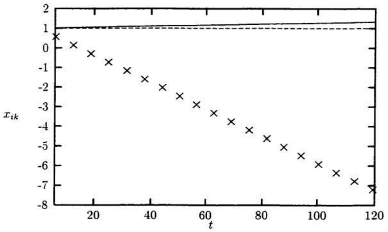

Figure 5.1: Numerical results for (5.9) with $c_1 = 1/2$. $k = 4$ ($\times$ $\times$), $k = 5$ (—), $k = 10$ ($- - -$)

5.2.2 Linear Case with $a(t) = -1$

If we take $a(t) = -1$ in (5.3), then clearly $\nu_1 = \int_0^T a(s)ds < 0$ and the differential equation has a unique, asymptotically stable periodic solution. If $a(t)$ is a negative constant, the differential equation could be scaled such that $a(t) = -1$. Then, we have the following theorem.

Theorem 5.2.2 *Suppose a linearized one-point collocation method is used to* *solve the linear nonautonomous differential equation (5.3) with $a(t) = -1$. Then, for fixed c_1 and k, the method admits a unique periodic solution that is asymptotically stable, provided*

$$k > \frac{T}{2}(1 - 2c_1).$$

Proof: If $a(t) = -1$, (5.4) simplifies to

$$x_{n+1} = x_n + \frac{T}{k + Tc_1}[-x_n + b(nT/k) + \frac{Tc_1}{k} b'(nT/k)]$$

or

$$x_{n+1} = x_n \left(\frac{k + T(c_1 - 1)}{k + Tc_1} \right) + \frac{T}{k + Tc_1}[b(nT/k) + \frac{Tc_1}{k} b'(nT/k)].$$

Proceeding in a manner analogous to the above, we can show by induction that

$$x_{n+1} = x_0 \left(\frac{k + T(c_1 - 1)}{k + Tc_1} \right)^{n+1} + \frac{T}{k + Tc_1} \sum_{r=0}^{n} \left(\frac{k + T(c_1 - 1)}{k + Tc_1} \right)^r \tilde{b}_{n-r} \tag{5.12}$$

where each $\tilde{b}_i$ is defined by (5.5).

Denoting the right-hand side of (5.12) by $\Pi(x_n)$, we can perform stroboscopic sampling and consider

$$x_k = \Pi^k(x_0) = x_0 \left(\frac{k + T(c_1 - 1)}{k + Tc_1} \right)^k + \frac{T}{k + Tc_1} \sum_{r=0}^{k-1} \left(\frac{k + T(c_1 - 1)}{k + Tc_1} \right)^r \tilde{b}_{k-r-1} \tag{5.13}$$

and the associated map

$$X_{n+1} = X_n \left(\frac{k+T(c_1-1)}{k+Tc_1}\right)^k + \frac{T}{k+Tc_1}\sum_{r=0}^{k-1}\left(\frac{k+T(c_1-1)}{k+Tc_1}\right)^r \tilde{b}_{k-r-1}. \tag{5.14}$$

Equation (5.14) is just the linear map

$$x_{n+1} = cX_n + d$$

with

$$c = \left(\frac{k+T(c_1-1)}{k+Tc_1}\right)^k$$

and

$$d = \frac{T}{k+Tc_1}\sum_{r=0}^{k-1}\left(\frac{k+T(c_1-1)}{k+Tc_1}\right)^r \tilde{b}_{k-r-1}.$$

The map has a single fixed point,

$$X^* = \frac{d}{1-c}, \tag{5.15}$$

which is asymptotically stable if and only if $|c| < 1$, that is,

$$k > \frac{T(1-2c_1)}{2},$$

and the result is established. □

The following results are simple consequences of the above theorem.

Theorem 5.2.3 *A linearized one-point collocation method with* $c_1 \in [\frac{1}{2}, 1]$, *applied to the linear nonautonomous differential equation with* $a(t) = -1$, *admits a unique periodic solution that is asymptotically stable for all values of* k.

Thus, the linearized implicit Euler method and the linearized implicit midpoint method for the linear nonautonomous differential equation with $a(t) = -1$, admit a unique, asymptotically stable periodic solution for all $k > 0$.

However,

Theorem 5.2.4 *The explicit Euler method for the linear nonautonomous differential equation with* $a(t) = -1$ *admits a unique, asymptotically stable periodic solution if and only if* $k > T/2$.

We attempt to bound $|X^*|$. The following lemma gives a bound on the solution $\Phi(t)$ of (5.3).

Lemma 5.2.5 *Let* $b(t)$ *be a* C^1 T*-periodic function of* t *and assume (5.3) with* $a(t) = -1$ *has a unique* T*-periodic solution. There exists a number* $M > 0$ *such that* $|b(t)|, |b'(t)| \leq M$ *and the* T*-periodic solution,* $\Phi(t)$*, satisfies*

$$|\Phi(t)| \leq M$$

for $t \to \infty$.

Proof: The boundedness of $b(t)$ and $b'(t)$ follows from the periodicity and continuity of both functions.

Let $\phi(t)$ be a periodic solution of (5.3) with $a(t) = -1$. Then, $\phi(t)$ satisfies the inequality

$$-\phi - M \leq \phi' \leq -\phi + M \qquad \text{for all t.}$$

This gives Gronwall's inequality

$$e^{-t}(x_0 + M) - M \leq \phi \leq e^{-t}(x_0 - M) + M, \tag{5.16}$$

see [16]. Thus, $\phi(t)$ is bounded for $t \geq 0$; therefore it approaches a T-periodic solution $\Phi(t)$. Taking $t \to \infty$ in (5.16) establishes the lemma. □

From the above lemma, we have $|X^*| \leq M$. Then, we have the following theorem.

Theorem 5.2.6 *Let X^*, given by (5.15), denote the fixed point generated by the stroboscopic sampling of the numerical solution of the linear nonautonomous differential equation (5.3) with $a(t) = -1$ by a linearized one-point collocation method. Then, the inequality*

$$|X^*| \leq M(1 + \frac{Tc_1}{k})$$

holds.

Proof: Since $b(t)$ is continuous, there is a number M such that $|b(t)|, |b'(t)| \leq M$ for all $t \in \mathbb{R}$. Then, for large t, the solution $|\Phi(t)| < M$.

Observe that

$$|\tilde{b}_r| \leq M_1 = M(1 + c_1 \frac{T}{k})$$

and

$$\begin{aligned}
|d| &\leq \frac{T}{k+Tc_1} \cdot M_1 \cdot \sum_{r=0}^{k-1} \left(\frac{k+T(c_1-1)}{k+Tc_1}\right)^r \\
&= \frac{T}{k+Tc_1} \cdot M_1 \cdot \left\{1 + \frac{k+T(c_1-1)}{T}[1 - \left(\frac{k+T(c_1-1)}{k+Tc_1}\right)^{k-1}]\right\} \\
&= M_1 \cdot [1 - \left(\frac{k+T(c_1-1)}{k+Tc_1}\right)^k] \\
&= M_1(1-c).
\end{aligned}$$

Therefore,

$$|X^*| \leq \frac{M_1(1-c)}{1-c} = M_1 = M(1 + \frac{Tc_1}{k}). \tag{5.17}$$

□

Hence, X^* has essentially the same bound as the periodic solution.

Numerical Experiments

Consider the linear nonautonomous equation

$$x' = -x + \sin t, \quad x(0) = 0 \tag{5.18}$$

which has a solution $x(t) = \frac{1}{2}[\sin t - \cos t + e^{-t}]$. Figure 5.2 shows the numerical results (stroboscopic sampling) of the linearized implicit midpoint method ($c_1 = 1/2$) with $k = 3$ and $k = 10$; Figure 5.3 shows the results of the explicit Euler method ($c_1 = 0$) with $k = 3$ and $k = 10$. In these experiments, $T = 2\pi$; hence, convergence to a unique periodic solution is expected for $k > \pi(1 - 2c_1)$.

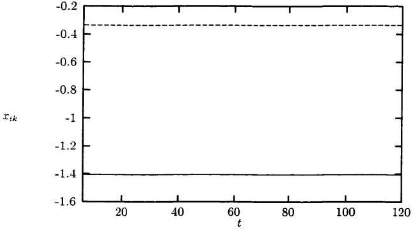

Figure 5.2: Numerical results of (5.18) using linearized implicit midpoint method: $k = 3$ (—) and $k = 10$ (− − −)

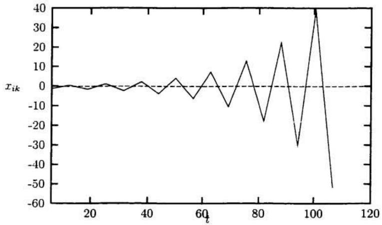

Figure 5.3: Numerical results of (5.18) using Explicit Euler method: $k = 3$ (—) and $k = 10$ ($- - -$)

5.2.3 Linear Case with $a(t) = -1 + \epsilon\rho(t)$

Let

$$a(t) = -1 + \epsilon\rho(t) \tag{5.19}$$

where $\epsilon \geq 0$ is a constant, and $\rho(t)$ is a T-periodic function of t. We assume that $-T + \epsilon \int_0^T \rho(s)ds < 0$, so that (5.3) has a unique asymptotically stable periodic solution.

Linearized One-Point Collocation Methods

We establish conditions under which a one-point collocation method with fixed c_1 and step-size $h > 0$ exhibits the same dynamical behaviour as the nonautonomous linear differential equation with $a(t) = -1 + \epsilon\rho(t)$.

Notation: In what follows, we will denote $b_n = b(nh)$, $b'_n = b'(nh)$, $\rho_n = \rho(nh)$ and $\rho'_n = \rho(nh)$. Here, as before, $T = hk$ $(k \in \mathbf{N})$.

Theorem 5.2.7 *Suppose a linearized one-point collocation method is used to solve a linear nonautonomous differential equation of the form (5.3) with* $a(t) = -1 + \epsilon\rho(t)$. *Then, for fixed* c_1 *and* k, *the method will admit a unique periodic solution that is asymptotically stable, provided*

$$\left|\prod_{i=0}^{k-1}[1 + \frac{kT(-1+\epsilon\rho_i) + c_1 T^2 \epsilon\rho'_i}{k^2 - kTc_1(-1+\epsilon\rho_i)}]\right| < 1.$$

Proof: From (5.4), we deduce that the linearized collocation methods, applied to the nonautonomous linear ODE with $a(t)$ given by (5.19), are

$$\begin{aligned} x_{n+1} &= \left[1 + \frac{h(-1+\epsilon\rho_n) + c_1 h^2 \epsilon\rho'_n}{1 - hc_1(-1+\epsilon\rho_n)}\right] x_n + \\ &\quad \frac{h}{1 - hc_1(-1+\epsilon\rho_n)} \cdot [b_n + c_1 h b'_n]. \end{aligned} \tag{5.20}$$

We can write the above as

$$x_{n+1} = T_n x_n + H_n \bar{b}_n := \Pi(x_n) \tag{5.21}$$

where

$$T_n = 1 + \frac{h(-1+\epsilon\rho_n) + c_1 h^2 \epsilon \rho_n'}{1 - hc_1(-1+\epsilon\rho_n)} \tag{5.22}$$

$$H_n = \frac{h}{1 - hc_1(-1+\epsilon\rho_n)} \tag{5.23}$$

$$\bar{b}_n = b_n + c_1 h b_n'. \tag{5.24}$$

Proceeding by induction, we establish that

$$x_{n+1} = x_0 \prod_{i=0}^{n} T_i + \sum_{i=0}^{n} H_i \bar{b}_i \prod_{j=i+1}^{n} T_j, \tag{5.25}$$

from which we deduce that

$$x_k = x_0 \prod_{i=0}^{k-1} T_i + \sum_{i=0}^{k-1} H_i \bar{b}_i \prod_{j=i+1}^{k-1} T_j := \Pi^k(x_0). \tag{5.26}$$

The discrete system that corresponds to stroboscopic sampling is the linear system

$$X_{n+1} = c(\epsilon) X_n + d(\epsilon), \tag{5.27}$$

where

$$c(\epsilon) = \prod_{i=0}^{k-1} T_i$$

$$d(\epsilon) = \sum_{i=0}^{k-1} H_i \bar{b}_i \prod_{j=i+1}^{k-1} T_j.$$

This system has a unique fixed point, X^*, given by (5.15). It is asymptotically stable if and only if $|c(\epsilon)| < 1$, that is, $|\prod_{i=0}^{k-1} T_i| < 1$, or

$$\left|\prod_{i=0}^{k-1}[1 + \frac{h(-1+\epsilon\rho_i) + c_1 h^2 \epsilon\rho_i'}{1 - hc_1(-1+\epsilon\rho_i)}]\right| < 1. \tag{5.28}$$

Substituting $h = T/k$ gives the result. □.

Numerical Experiments

For each of the three special values of c_1, $\rho(t) = \sin t$ and increasing values of ϵ, we determined, using (5.28), the minimum value of k such that each method has dynamical behaviour that is the same as that of the differential equation. The results are illustrated in Figure 5.4.

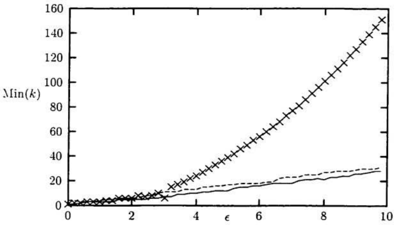

Figure 5.4: Least k for unique periodic solution. Explicit Euler (– – –), Linearized Implicit Midpoint (—), Linearized Implicit Euler (× ×)

For $\epsilon < 2$, the Explicit Euler method is the most restrictive of the three (that is, comparatively larger minimum values of k must be used to obtain dynamical behaviour that is the same as that of the differential equation). However, as ϵ is increased, the Explicit Euler method outperforms the linearized implicit Euler method by becoming less restrictive than that method for $\epsilon \geq 3$. For $\epsilon \geq 3$, the explicit Euler and linearized implicit midpoint methods give comparable results, and for those values of ϵ, the linearized implicit Euler method becomes more and more restrictive in comparison to the other two.

Finally, we develop a bound on $|X^*|$. For comparison purposes, we present the following lemma, which can be proved in a manner analogous to Lemma 5.2.5.

Lemma 5.2.8 *Let $b(t), \rho(t)$ be C^1 T-periodic functions of t. There exist numbers $M, M_2 > 0$ such that $|b(t)|, |b'(t)| \leq M$, $|\rho(t)|, |\rho'(t)| \leq M_2$, and the solution $\Phi(t)$ satisfies*

$$|\Phi(t)| \leq \frac{M}{|1 - \epsilon M_2|}$$

as $t \to \infty$.

Theorem 5.2.9 *Let X^*, given by (5.15), be the fixed point generated by the stroboscopic sampling of the numerical solution of the linear nonautonomous differential equation with $a(t) = -1 + \epsilon\rho(t)$ by a linearized collocation method. Then, the following inequality holds.*

$$|X^*| \leq \frac{M(1 + \frac{Tc_1}{k})}{|1 - c_1\epsilon M_2 - \epsilon M_2|}. \tag{5.29}$$

Proof: Since $b(t)$ and $\rho(t)$ are C^1 and periodic, there are numbers M, M_2 such that $|b(t)|, |b'(t)| \leq M$ and $|\rho(t)|, |\rho'(t)| \leq M_2$ for all $t \in \mathbb{R}$.

For each i,

$$\begin{aligned}
|T_i| &\leq \left|\frac{1 + hc_1 + h\epsilon M_2 - h}{1 + hc_1 - hc_1\epsilon M_2}\right| \\
|H_i| &\leq \frac{h}{|1 + hc_1 - hc_1\epsilon M_2|} \\
|\bar{b}_i| &\leq M_1 = M(1 + c_1 h),
\end{aligned}$$

where $h = T/k$. If

$$v(\epsilon) = \frac{1 + hc_1 - hc_1\epsilon M_2 + h\epsilon M_2 + h^2 c_1 \epsilon M_2 - h}{1 + hc_1 - hc_1\epsilon M_2}, \tag{5.30}$$

then

$$\begin{aligned}
|d(\epsilon)| &\leq \frac{h}{|1 + hc_1 - hc_1\epsilon M_2|} \cdot M_1 \cdot \left|\sum_{i=0}^{k-1} v^{k-i-1}\right| \\
&= \frac{hM_1}{|1 + hc_1 - hc_1\epsilon M_2|} \cdot \left|\frac{1 - v^k}{h(1 - c_1 h\epsilon M_2 - \epsilon M_2)}\right| \cdot |1 + hc_1 - hc_1\epsilon M_2| \\
&= M_1 \cdot \left|\frac{1 - v^k}{(1 - c_1 h\epsilon M_2 - \epsilon M_2)}\right|.
\end{aligned}$$

On the other hand, we deduce from the definition of $c(\epsilon)$ that

$$|1 - c(\epsilon)| \geq \left|1 - \left(\frac{1 + hc_1 - h + h\epsilon M_2 - hc_1\epsilon M_2 + h^2 c_1 \epsilon M_2}{1 + hc_1 - hc_1\epsilon M_2}\right)^k\right|.$$

Hence,

$$|X^*(\epsilon)| \leq \frac{M_1}{|1 - c_1 h\epsilon M_2 - \epsilon M_2|} = \frac{M(1 + c_1 h)}{|1 - c_1 h\epsilon M_2 - \epsilon M_2|}, \tag{5.31}$$

which is identical to the inequality (5.29). □

If we substitute $\epsilon = 0$ in (5.31), we obtain (5.17) as expected. Here, as well, the bound for $X^*(\epsilon)$ is the same as that of the periodic solution as $h \to 0$.

We would like to obtain a relationship between the dynamical approach study and stability analysis. We introduce a natural stability criterion for the differential equation as well as any numerical method used to discretize it.

5.2.4 Conditional AN-stability and AN-stability

We consider the problem of determining a criterion for some sort of "controlled behaviour" of the solutions of the methods. We adopt a linear stability criterion that is based on the scalar test equation

$$x' = a(t)x, \tag{5.32}$$

where $a(t) \in \mathbb{C}$. If $\text{Re}(a(t)) < 0$ for all $t \in [\beta_1, \beta_2]$, then

$$x(t+h) = Kx(t), \qquad |K| \leq 1,$$

for all $x \in [\beta_1, \beta_2]$ and $h > 0$.

Definition 5.2.10 *A numerical method is said to be conditionally* **AN**-*stable for some $h > 0$ if, when applied to the test equation (5.32) with $Re(a(t)) < 0$*

for all t,

$$x_{n+1} = \tilde{K}(h)x_n, \qquad |\tilde{K}(h)| \leq 1,$$

holds for all $n \in \mathbf{N}$.

If this condition is satisfied for all $h > 0$, then the method is **AN**-*stable (Lambert [23], Stuart & Humphries [31]).*

The following simple result gives a condition under which the linearized one-point collocation methods are conditionally **AN**-stable.

Theorem 5.2.11 *The linearized one-point collocation methods are conditionally* **AN**-*stable if*

$$(1 - 2c_1)a(t) + c_1 h a'(t) \geq -\frac{2}{h} \tag{5.33}$$

and $a(t) + c_1 h a'(t) \leq 0$.

Proof: Use the above test equation in (1.23).

Examples:

1. If $a(t) = -1$, then the linearized one-point collocation methods are conditionally **AN**-stable if

$$\frac{h}{2}(1 - 2c_1) \leq 1.$$

It is easy to observe that the methods are **AN**-stable if $c_1 \geq \frac{1}{2}$ and conditionally **AN**-stable for sufficiently small h if $c_1 < \frac{1}{2}$.

2. If $a(t) = -1 + \epsilon\rho(t)$, where ϵ, $\rho(t) \in \mathbf{C}$, and $\epsilon(\rho(t) + c_1 h\rho'(t)) \leq 1$, then the linearized one-point collocation methods are conditionally **AN**-stable if

$$-1 + \epsilon\rho(t) + c_1(h\epsilon\rho'(t) + 2 - 2\epsilon\rho(t) \geq -\frac{2}{h}.$$

Our results suggest the existence of a relationship between the linear stability theory of the collocation methods and the existence and asymptotic stability of periodic solutions, identified via stroboscopic sampling. This relationship is stated in the following theorem.

Theorem 5.2.12 *Suppose a linearized one-point collocation method is used to solve a linear nonautonomous equation with periodic coefficients which has a unique, asymptotically stable periodic solution. Then, the method yields the same dynamical behaviour if it is conditionally* **AN***-stable. The reverse does not necessarily hold.*

Proof: The first statement follows from the conclusions of Theorems 5.2.2 and 5.2.7 as well as Examples 1 and 2 in section 5.2.4.

We use an example to show that the reverse is not true in general. If, in § 5.2.3, we let $\rho(t) = \sin t$, $c_1 = 1$, $k = 10$, and $\epsilon = 2$, we find that the method has a single asymptotically stable periodic solution since it satisfies the condition (5.28). However, it is not conditionally **AN**-stable since the condition in Example 2 is violated. □

The last theorem is very significant, since it gives us a bridge connecting standard stability theory with dynamical systems. Naturally, we would like to find out if there is a corresponding result for the nonlinear case, which we now consider.

5.3 Nonlinear Case

We consider the nonlinear equation,

$$x' = a(t)g(x) + b(t), \qquad \text{for } t \geq 0 \text{ and } x(0) = x_0 \tag{5.34}$$

where $b(t)$ is a T-periodic function of t and $g(x)$ is a C^2 nonlinear function of x.

The following is an existence and uniqueness theorem for the solution of (5.34).

Theorem 5.3.1 *The differential equation (5.34) has a unique T-periodic solution that is asymptotically stable if $g(x) \in C^2$, $g'(x) > 0$, $a(t) < 0$ for all x and t, and $g(x) \to +\infty$ as $x \to +\infty$, $g(x) \to -\infty$ as $x \to -\infty$.*

Proof: We follow a method of proof similar to that used in Hale & Koçac [16]. We start by establishing the existence of at least one periodic solution of (5.34). Let

$$\begin{aligned} a_{\min} &:= \min_{t\geq 0} a(t) \\ a_{\max} &:= \max_{t\geq 0} a(t) \\ b_{\min} &:= \min_{t\geq 0} b(t) \\ b_{\max} &:= \max_{t\geq 0} b(t), \end{aligned}$$

where $a(t)$ and $b(t)$ are T-periodic.

If $\Phi(t)$ is any solution of (5.34), then $\Phi(t)$ satisfies

$$g(\Phi)a_{\min} + b_{\min} \leq \Phi' \leq g(\Phi)a_{\max} + b_{\max}. \tag{5.35}$$

Defining the sets

$$\begin{aligned} U_- &= \{(t,x) : g(x)a_{\min} + b_{\min} > 0\} \\ U_+ &= \{(t,x) : g(x)a_{\max} + b_{\max} < 0\} \end{aligned}$$

we observe that Φ is increasing in U_- and decreasing in U_+. If $g(x) \to +\infty$ as $x \to +\infty$ then $x \in U_-$ for all x sufficiently large, and if $g(x) \to -\infty$ as $x \to -\infty$ then $x \in U_+$ for all x sufficiently small.

Therefore, the solution is bounded for all $t \geq 0$; hence, there is at least one T-periodic solution of (5.34).

Uniqueness: Let $x_1(t)$ and $x_2(t)$ be two T-periodic solutions of the differential equation. Then,

$$x_1' = a(t)g(x_1) + b(t) \tag{5.36}$$

$$x_2' = a(t)g(x_2) + b(t). \tag{5.37}$$

From (5.36), (5.37) and the Mean Value Theorem, we have

$$X'(t) = a(t)g'(\alpha(t))X \tag{5.38}$$

where $X(t) = x_1(t) - x_2(t)$ and $\alpha(t)$ is between $x_1(t)$ and $x_2(t)$ for all t.

If $g'(x) > 0$ for all x, and $a(t) \leq -\epsilon < 0$ for all t, then clearly $X(t) \to 0$ as $t \to \infty$, proving the uniqueness of the periodic solution. □

We assume the hypotheses in Theorem 5.3.1 are satisfied, and consider the exercise of devising numerical approximations for the solution of (5.34).

5.3.1 Linearized One-Point Collocation Methods

The linearized collocation methods, applied to (5.34), are given by

$$x_{n+1} = x_n + \frac{h\tilde{a}_n g(x_n)}{1 - ha_n c_1 g'(x_n)}h + \frac{h\tilde{b}_n}{1 - ha_n c_1 g'(x_n)} \tag{5.39}$$

where $\tilde{b}_n = b_n + c_1 hb_n'$ and $\tilde{a}_n = a_n + c_1 ha_n'$.

We perform a simplification on the third term of (5.39) that takes the form of evaluating the derivative of g at the starting value, instead at each step. The resulting method, that will be referred to as a simplified linearized one-point collocation method, is

$$x_{n+1} = x_n + h\tilde{a}_n G(x_n; n) + H_n \tilde{b}_n, \tag{5.40}$$

where

$$\begin{aligned} H_n &= \frac{h}{1 - ha_n c_1 g'(x_0)} \\ G(x; n) &= \frac{g(x)}{1 - ha_n c_1 g'(x)}. \end{aligned}$$

5.3.2 The Dynamical Systems Approach

We would like to take the dynamical systems approach and determine the conditions under which (5.40), applied to the differential equation (5.34), yields the same dynamics as the continuous system.

We rewrite (5.40) as

$$h\tilde{a}_n G(x_n; n) = x_{n+1} - x_n - H_n \tilde{b}_n.$$

Inductively, we can show that

$$x_{n+1} = x_0 + \sum_{r=0}^{n} H_r \tilde{b}_r + h \sum_{r=0}^{n} \tilde{a}_r G(x_r; r). \tag{5.41}$$

We choose an integer k such that $T = hk$. Sampling stroboscopically in the iteration above, we get

$$x_k = x_0 + \sum_{r=0}^{k-1} H_r \tilde{b}_r + h \sum_{r=0}^{k-1} \tilde{a}_r G(x_r; r) \tag{5.42}$$

and associate this with the discrete system

$$X_{n+1} = X_n + \sum_{r=0}^{k-1} H_r \tilde{b}_r + h \sum_{r=0}^{k-1} \tilde{a}_r G(\tilde{X}_r; r) \tag{5.43}$$

where,

$$\begin{aligned} \tilde{X}_0 &= X_n \\ \tilde{X}_r &= X_n + \sum_{i=0}^{r-1} H_r \tilde{b}_r + h \sum_{i=0}^{r-1} \tilde{a}_i G(\tilde{X}_i; i) \text{ for } r = 1, 2, \ldots, k-1. \end{aligned} \tag{5.44}$$

The fixed points of (5.43) correspond to periodic solutions of (5.34). Fixed points are points, X^*, such that

$$\sum_{r=0}^{k-1} H_r \tilde{b}_r + h \sum_{r=0}^{k-1} \tilde{a}_r G(\tilde{X}_r; r) = 0, \tag{5.45}$$

where

$$\tilde{X}_r = X^* + \sum_{i=0}^{r-1} H_i \tilde{b}_i + h \sum_{i=1}^{r-1} \tilde{a}_i G(\tilde{X}_i; i) \text{ for } r = 1, 2, \ldots, k-1,$$

and $X_0 = X^*$. Define the sequence of functions $F_1(x)$, $F_2(x)$, ... by

$$F_k(x) = \sum_{r=0}^{k-1} [\tilde{a}_r h G(\tilde{x}_r; r) + H_r \tilde{b}_r] \tag{5.46}$$

where $\tilde{x}_0 = x$ and

$$\tilde{x}_r = x + \sum_{i=0}^{r-1} [\tilde{a}_i h G(\tilde{x}_i; i) + H_i \tilde{b}_i] \tag{5.47}$$

for $r = 1, 2, \ldots, k-1$.

The theorem below gives conditions under which the simplified linearized one-point collocation methods, applied to (5.34), exhibit dynamical behaviour that is the same as the differential equation.

Lemma 5.3.2 *If* $|1 + h\tilde{a}(t)G'(x;t)| < 1$ *for all* x *and* t*, then*

$$-1 \leq F_k'(x) < 0$$

for all $k \in \mathbf{N}$*, where* $F_k(x)$ *is given by (5.46).*

Proof: Observe that $F_1(x) = h\tilde{a}_0 G(x;0) + H_0\tilde{b}_0$ and

$$F_k(x) = F_{k-1}(x) + h\tilde{a}_{k-1}G(x + F_{k-1}(x); k-1) + H_{k-1}\tilde{b}_{k-1} \tag{5.48}$$

for $k = 2, 3, \ldots$. It can be established by induction that

$$F_k'(x) = F_1'(x)\prod_{i=1}^{k-1}(1+\overline{a}_i) + \sum_{i=1}^{k-1}\left[\overline{a}_i \prod_{j=i+1}^{k-1}(1+\overline{a}_j)\right] \tag{5.49}$$

where $\overline{a}_i = h\tilde{a}_i G'(x + F_i(x); i) < 0$ for each $i \in \mathbf{N}$.

It can be shown by induction on k that $F_k'(x) < 0$ for all x and $k \in \mathbf{N}$. On the other hand, we obtain an extremum, $\overline{F}_k'$, of F_k' when

$$\frac{\partial F'}{\partial a_l} = F_1'(x)\prod_{i=1, i\neq l}^{k-1}(1+\overline{a}_i) + \sum_{i=1}^{l-1}\left[\overline{a}_i \prod_{j=i+1, j\neq l}^{k-1}(1+\overline{a}_j)\right] + \prod_{j=l+1}^{k-1}(1+\overline{a}_j) = 0$$

for $l = 1, 2, \ldots, k-1$. Taking $l = k-1$ in the above equation gives

$$F_1'(x)\prod_{i=1}^{k-2}(1+\overline{a}_i) + \sum_{i=1}^{k-2}\left[\overline{a}_i \prod_{j=i+1}^{k-2}(1+\overline{a}_j)\right] = -1, \tag{5.50}$$

from which we deduce that $\overline{F}'_k = -1$. Clearly, this extremum is a minimum, and the lemma is established. □

Theorem 5.3.3 *Assume the differential equation (5.34) has a unique solution. If*

(i) $g''(x)g(x) \leq 0$,

(ii) $c_1 a'(t) < \frac{-a(t)}{h}$ *for all* t, *and*

(iii) Max $G'(x;t) \leq \frac{1}{h|\bar{a}(t)|}$,

then a simplified linearized one-point collocation method has a periodic solution for any $k = T/h$*; this solution is unique and asymptotically stable.*

Proof: Finding possible fixed points of (5.43), hence periodic solutions associated with the methods, is the same as finding zeros of (5.46). This is equivalent to solving the nonlinear system:

$$
\begin{aligned}
-h\tilde{a}_0 G(x^*;0) - h\tilde{a}_1 G(\tilde{x_1};1) - h\tilde{a}_2 G(\tilde{x_2};2) - \dots & \\
-h\tilde{a}_{k-1} G(\tilde{x_{k-1}};k-1) - H_0\tilde{b}_0 - H_1\tilde{b}_1 - \dots - H_{k-1}\tilde{b}_{k-1} &= 0 \\
\tilde{x_1} - x^* - h\tilde{a}_0 G(x^*;0) - H_0\tilde{b}_0 &= 0 \\
\tilde{x_2} - x^* - h\tilde{a}_0 G(x^*;0) - h\tilde{a}_1 G(\tilde{x_1};1) - H_0\tilde{b}_0 - H_1\tilde{b}_1 &= 0 \\
&\vdots \\
\tilde{x_{k-1}} - x^* - h\tilde{a}_0 G(x^*;0) - h\tilde{a}_1 G(\tilde{x_1};1) - h\tilde{a}_2 G(\tilde{x_2};2) - \dots & \\
-h\tilde{a}_{k-2} G(\tilde{x_{k-2}};k-2) - H_0\tilde{b}_0 - H_1\tilde{b}_1 - \dots - H_{k-2}\tilde{b}_{k-2} &= 0.
\end{aligned}
$$

The above system is of the form $\mathbf{F(x)} = \mathbf{0}$, where $\mathbf{F}$ is a nonlinear function of

$$
\mathbf{x} = \begin{pmatrix} x^* = \tilde{x_0} \\ \tilde{x_1} \\ \tilde{x_2} \\ \vdots \\ \tilde{x_{k-1}} \end{pmatrix}
$$

.

To prove existence, it is sufficient to show that the Jacobian matrix, $\mathrm{J}(\mathbf{F})$, of $\mathbf{F}$ is non-singular. Now,

$$
\mathrm{J(F)} = \begin{pmatrix}
d_0 & d_1 & d_2 & d_3 & \dots & d_{k-2} & d_{k-1} \\
-1+d_0 & 1 & 0 & 0 & \dots & 0 & 0 \\
-1+d_0 & d_1 & 1 & 0 & 0 & \dots & 0 \\
-1+d_0 & d_1 & d_2 & 1 & 0 & \dots & 0 \\
 & & & \vdots & & & \\
-1+d_0 & d_1 & d_2 & \dots & d_{k-2} & d_{k-1} & 1
\end{pmatrix}
$$

where $d_i = -h\tilde{a}_i G'(\bar{x}_i; i)$ for each i. We perform one elementary row operation: row 1 $\rightarrow$ row 1 - row k. The matrix becomes

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 & -1+d_{k-1} \\ -1+d_0 & 1 & 0 & 0 & \dots & 0 & 0 \\ -1+d_0 & d_1 & 1 & 0 & 0 & \dots & 0 \\ -1+d_0 & d_1 & d_2 & 1 & 0 & \dots & 0 \\ & & & \vdots & & & \\ -1+d_0 & d_1 & d_2 & \dots & d_{k-2} & d_{k-1} & 1 \end{pmatrix}$$

It is easy to see that the above matrix is nonsingular.

To prove uniqueness, let x_1 and x_2 be two fixed points of (5.43). Then, from (5.48),

$$\begin{aligned} F_k(x_1) &= F_{k-1}(x_1) + h\tilde{a}_{k-1}G(x_1 + F_{k-1}(x_1); k-1) \\ &\quad + H_{k-1}\bar{b}_{k-1} = 0 \end{aligned} \tag{5.51}$$

and

$$\begin{aligned} F_k(x_2) &= F_{k-1}(x_2) + h\tilde{a}_{k-1}G(x_2 + F_{k-1}(x_2); k-1) \\ &\quad + H_{k-1}\bar{b}_{k-1} = 0. \end{aligned} \tag{5.52}$$

Subtracting (5.52) from (5.51) and using the Mean Value Theorem gives

$$(x_1 - x_2)\{F'_{k-1}(\alpha_1) + h\tilde{a}_{k-1}G'(\alpha_2; k-1)(1 + F'_{k-1}(\alpha_1))\} = 0, \tag{5.53}$$

which may alternatively be written as

$$(x_1 - x_2)\{F'_{k-1}(\alpha_1)(1 + h\tilde{a}_{k-1}G'(\alpha_2; k-1)) + h\tilde{a}_{k-1}G'(\alpha_2; k-1)\} = 0. \tag{5.54}$$

In the above equations, α_1 is between x_1 and x_2 and α_2 is between $x_1 + F_{k-1}(x_1)$ and $x_2 + F_{k-1}(x_2)$.

From the hypotheses of the theorem, we have that $F' < 0$, $G' > 0$, $\tilde{a}_{k-1} < 0$ and $1 \geq 1 + h\tilde{a}_{k-1}G'(\alpha_2; k-1) \geq 0$. Thus, $x_1 = x_2$.

To prove that the periodic solution is asymptotically stable, we need only to show that $|1 + F'| < 1$. Since $1 \geq 1 + h\tilde{a}(t)G'(x; t) \geq 0$, we have $-1 \leq F' < 0$ from Lemma 5.3.2.

This completes the proof. □

5.3.3 Nonlinear Stability Theory

We wish to establish conditions under which numerical methods for the solution of (5.34) behave in a "controlled" manner.

Consider $\mathbb{R}^m$ as an inner product space with corresponding norm $\|\cdot\|$. Then, if neighbouring solution curves converge with respect to this norm (as $t \to \infty$), the system is said to have **contractive** solutions (Lambert [23], Stuart & Humphries [31]). As in the linear case, we will contrast the conditions for stability with the existence and uniqueness of a periodic solution in the numerical methods.

We briefly state the concepts of contractivity and conditional **BN**-stability.

Definition 5.3.4 *Let $x(t)$ and $\overline{x}(t)$ be any two solutions of the differential equation $x' = f(x, t)$, satisfying initial conditions $x(0) = \eta$, $\overline{x}(0) = \overline{\eta}$, $\eta \neq \overline{\eta}$. Then, if*

$$\|x(t_2) - \overline{x}(t_2)\| \leq \|x(t_1) - \overline{x}(t_1)\|$$

holds under the $\mathbb{R}^m$ norm $\|\cdot\|$ for all t_1, t_2 such that $\beta_1 \leq t_1 \leq t_2 \leq \beta_2$, the solutions of the system are said to be **contractive** *in $[\beta_1, \beta_2]$.*

The discrete analog of the above definition is given below.

Definition 5.3.5 *Let $\{x_n\}$ and $\{\overline{x}_n\}$ be two numerical solutions generated by a numerical method with different starting values. Then, if*

$$||x_{n+1} - \overline{x}_{n+1}|| \leq ||x_n - \overline{x}_n||, \qquad 0 \leq n \leq N,$$

the numerical solutions are said to be contractive for $n \in [0, N]$.

Definition 5.3.6 *The system $x' = f(x, t)$ is* **dissipative** *in $[\beta_1, \beta_2]$ if*

$$< f(x, t) - f(\overline{x}, t), x - \overline{x} > \leq 0 \tag{5.55}$$

holds for all x, $\overline{x} \in \mathbb{R}^m$ and for all $t \in [\beta_1, \beta_2]$.

It is easy to show that the solutions of a dissipative system are contractive under the norm induced by the inner product in (5.55). It is desirable that a numerical method, used with fixed stepsize $h > 0$ to solve a dissipative system, gives contractive solutions. This brings us to the concepts of conditional **BN**-stability and **BN**-stability, nonlinear nonautonomous stability criteria.

Definition 5.3.7 *If a numerical method, applied with fixed steplength $h > 0$ to (5.34) satisfying (5.55), generates contractive solutions, the method is said to be conditionally* **BN**-*stable. If the method generates contractive solutions when applied with* any $h > 0$, *then it it is* **BN**-*stable (Lambert [23]).*

The concepts of **AN**- and **BN**-stability are equivalent for nonconfluent Runge-Kutta methods.

To determine the conditional **BN**-stability of the methods, we use the scalar test system

$$x' = a(t)g(x) \tag{5.56}$$

where, as before, $a(t) \in \mathbb{C}$ and $g(x) \in C^1$. This system is dissipative if

$$\begin{aligned} a(t) < g(x) - g(\overline{x}), x - \overline{x} > &= a(t)g'(\xi)|x - \overline{x}|^2 \\ &\leq 0, \end{aligned}$$

where ξ lies between x and $\overline{x}$ and $< \cdot, \cdot >$ is an inner product in $\mathbb{R}$. The condition is satisfied if $a(t) \leq 0$ for all t and $g'(x) \geq 0$ for all x. Therefore, the existence and uniqueness of a periodic solution in (5.34), by Theorem 5.3.1, is a sufficient condition for the dissipativity of the system.

Now, we determine the conditions under which the linearized one-point collocation methods are conditionally **BN**-stable. Applying the methods to the test system (5.56) for two different initial conditions, gives

$$x_{n+1} = x_n + h\tilde{a_n}G(x_n; n) \tag{5.57}$$

$$\overline{x}_{n+1} = \overline{x}_n + h\tilde{a_n}G(\overline{x}_n; n). \tag{5.58}$$

Subtracting (5.58) from (5.57) and using the Mean Value Theorem gives

$$|x_{n+1} - \overline{x}_{n+1}| = |(x_n - \overline{x}_n)| \cdot |1 + h\tilde{a_n}G'(\xi_n; n)|, \tag{5.59}$$

where ξ_n lies between x_n and $\overline{x}_n$.

The solutions generated by the methods are contractive if $|1+h\tilde{a_n}G'(\xi_n; n)| \leq 1$. Assuming $\tilde{a}(t) < 0$ for all t and $G'(x; t) \geq 0$ for all x and t, the methods are conditionally **BN**-stable if

$$G'(x; t) \leq \frac{2}{h|\tilde{a}(t)|}. \tag{5.60}$$

Remark: If we let $g(x) = x$ (the linear case), then condition (5.60) collapses to (5.33), which is the condition for conditional **AN**-stability.

5.3.4 Discussion

We have, in Theorem 5.3.3, established *sufficient* conditions for the simplified linearized one-point collocation methods to exhibit the same dynamics as the nonlinear nonautonomous ODE. Conditions (i) and (ii) are required for conditional **BN**-stability as well, but condition (iii) is not *necessary* for the existence of a unique, asymptotically stable solution. In fact, we can come to the same conclusion as in Theorem 5.3.3 if, in (5.53),

$$F'(x)[1 + h\tilde{a}(t)G'(x;t)] + h\tilde{a}(t)G'(x;t) \neq 0 \qquad (5.61)$$

for all x and t. The following theorem shows that condition (5.61) is satisfied if the method is conditionally **BN**-stable.

Theorem 5.3.8 *Suppose a simplified linearized one-point collocation method is used to solve a nonlinear nonautonomous equation of the form (5.34) with periodic coefficients which has a unique, asymptotically stable periodic solution. Then, the method yields a unique periodic solution that is asymptotically stable if it is conditionally* **BN***-stable.*

Proof: Assume conditional **BN**-stability. To prove the theorem, it is sufficient to establish condition (5.61). From conditional **BN**-stability, we have $-1 \leq 1 + h\tilde{a}G' \leq 1$. Recall that $-1 \leq F' < 0$ from Lemma 5.3.2.

Therefore,

$$F' + h\tilde{a}G' \leq F'(1 + h\tilde{a}G') + h\tilde{a}G' \leq -F' + h\tilde{a}G'. \tag{5.62}$$

From Lemma 5.3.2, $0 < -F' \leq 1$, so that if $-1 \leq 1 + h\tilde{a}G' < 0$,

$$F'(1 + h\tilde{a}G') + h\tilde{a}G' \leq 1 + h\tilde{a}G' < 0.$$

If $0 \leq 1 + h\tilde{a}G' \leq 1$,

$$F' + h\tilde{a}G' \leq F'(1 + h\tilde{a}G') + h\tilde{a}G' \leq h\tilde{a}G' < 0. \tag{5.63}$$

This proves the theorem. □

Remark: The above theorem is equivalent to Theorem 5.2.12 (in the linear case) if we replace the concept of conditional **BN**-stability by conditional **AN**-stability.

5.4 Conclusion

We already knew that numerical methods can introduce spurious behaviour into the solution for autonomous equations. Concentrating on linear and nonlinear nonautonomous equations with unique periodic solutions, and discretizing them using linearized one-point collocation methods, we were interested in the existence of periodic solutions in the numerical methods.

We found that the results obtained from the dynamical systems approach are closely linked to those that are imposed by standard stability analysis. It has been shown that, for linear and nonlinear nonautonomous differential equations of the form considered in this chapter, the conditional **AN**- or **BN**-stability of a linearized one-point collocation method is a sufficient condition for the method yielding the same dynamical behaviour as the differential equation under consideration.

Chapter 6

Higher Dimensional Linear Systems

6.1 Introduction

In this chapter, we will extend some of the results of the preceding one to higher dimensional linear systems. In particular, we would like to determine conditions under which the results that were obtained in the scalar case carry over.

Consider the system of linear equations (see Sanchez [27]),

$$x' = A(t)x + b(t), \tag{6.1}$$

where $A(t) = [a_{ij}(t)] \in \mathbb{R}^{m,m}$ is a continuous periodic matrix of period T and $b(t) = (b_i(t)) \in \mathbb{R}^m$ is also continuous and T-periodic. In addition, consider the boundary conditions

$$x(0) - x(T) = 0. \tag{6.2}$$

We will also be concerned with the corresponding homogeneous system

$$x' = A(t)x. \tag{6.3}$$

Since $A(t)$, $b(t)$ are T-periodic, then any solution $x(t)$ of (6.1) or (6.3) satisfying (6.2) is T-periodic (Hartman [17]). The theorem below is an existence and uniqueness result for the solution of (6.1).

Theorem 6.1.1 (Hartman [17]) *Let $A(t)$ be continuous for $0 \leq t \leq T$ and T-periodic. Then, (6.1) has a T-periodic solution $x(t)$ satisfying (6.2) for every continuous T-periodic $b(t)$ if and only if (6.3),(6.2) has no nontrivial ($\not\equiv 0$) solution; in which case $x(t)$ is unique and there exists a constant K, independent of $b(t)$, such that*

$$\|x(t)\| \leq K \int_0^T \|b(s)\| ds \quad \text{for } 0 \leq t \leq T. \tag{6.4}$$

We now consider linearized one-point collocation methods, as used to discretize a system of the form (6.1) which satisfies the hypotheses of Theorem 6.1.1. In a manner similar to the discussion of the linear scalar case of Chapter 5 with $a(t) = -1 + \epsilon\rho(t)$, we will use stroboscopic sampling to establish conditions under which the methods exhibit the same dynamics as the differential equation.

6.2 Linearized Collocation Methods

We rewrite the linearized one-point collocation methods (1.23) in the form

$$[I - c_1 h f_x(t_n, x_n)]x_{n+1} = [I - c_1 h f_x(t_n, x_n)]x_n + h[f(t_n, x_n) + c_1 h f_t(t_n, x_n)] \tag{6.5}$$

where, as before, I is the $m \times m$ identity matrix, h a constant stepsize and f_x the Jacobian of f with respect to x.

Applying (6.5) to (6.1), we obtain

$$[I - hc_1 A^{(n)}]x^{(n+1)} = [I - hc_1 A^{(n)}]x^{(n)} + h^2 c_1 [A^{(n)'} x^{(n)} + b^{(n)'}] + h(A^{(n)} x^{(n)} + b^{(n)})$$

where $A^{(n)} = [a_{ij}(nh)]$, $A^{(n)'} = [a'_{ij}(nh)]$, etc.

The methods can then be written in the form of the discrete system

$$x^{(n+1)} = S^{(n)} x^{(n)} + H^{(n)} \tilde{b}^{(n)} \tag{6.6}$$

where

$$\begin{aligned} S^{(n)} &= [I + h(I - hc_1 A^{(n)})^{-1} A^{(n)} + h^2 c_1 (I - hc_1 A^{(n)})^{-1} A^{(n)'}], \\ H^{(n)} &= h(I - hc_1 A^{(n)})^{-1}, \\ \tilde{b}^{(n)} &= b^{(n)} + hc_1 b^{(n)'}. \end{aligned}$$

It can be verified by induction that the system (6.6) can be written

$$x^{(n+1)} = \prod_{i=0}^{n} S^{(n-i)} x^{(0)} + \sum_{i=0}^{n} \left(\prod_{j=i+1}^{n} V^{(j)} \right) H^{(i)} \tilde{b}^{(i)} \tag{6.7}$$

where $V^{(j)} := S^{(n-j+i+1)}$.

6.2.1 Periodic Solutions

We now consider the possible existence of a unique, asymptotically stable periodic solution in the discrete system. For any $k \in \mathbf{N}$, we define the matrix

$$S = \prod_{i=0}^{k-1} S^{(k-1-i)}. \tag{6.8}$$

Then, we have the following theorem.

Theorem 6.2.1 *Suppose a linearized one-point collocation method is used to discretize the system (6.1) with boundary conditions (6.2), satisfying the hypotheses of Theorem 6.1.1. Then, the method has a unique, asymptotically stable periodic solution provided the matrix S, defined by (6.8), has all of its eigenvalues with moduli less than unity.*

Proof: From (6.7), we have

$$x^{(k)} = \prod_{i=0}^{k-1} S^{(i)} x^{(0)} + \sum_{i=0}^{k-1} \left(\prod_{j=i+1}^{k-1} V^{(j)} \right) H^{(i)} \tilde{b}^{(i)}. \tag{6.9}$$

Choose h such that $T = hk$, $k \in \mathbf{N}$. The discrete system corresponding to the stroboscopic sampling of (6.7) is then given by

$$x^{(n+1)} = Sx^{(n)} + v, \tag{6.10}$$

where S is defined by (6.8) and

$$v = \sum_{i=0}^{k-1} \left(\prod_{j=i+1}^{k-1} V^{(j)} \right) H^{(i)} \bar{b}^{(i)}.$$

Clearly, the system (6.10) has a unique fixed point given by

$$x^* = (I - S)^{-1} v$$

provided the matrix S has no eigenvalue of 1. This fixed point is asymptotically stable if all eigenvalues of S have moduli less than unity. □

Referring back to § 5.2.3, we observe that the argument has been generalized; the role of T_i in (5.22) is replaced by the matrices $S^{(i)}$ and the condition (5.28) is replaced by the conclusion of Theorem 6.2.1. A similar link exists between the vector v and the corresponding term in (5.25).

6.2.2 Remarks

The complexity of the matrix S and the vector v in the above argument shows the difficulties that arise in moving from the scalar case to higher dimensions. Future work will be directed at relating the conditions imposed by Theorem 6.2.1 on S to the original matrix $A(t)$. This will enable us to characterize those systems, if any, which could result in spurious periodic solutions in linearized one-point collocation methods.

The result on the link between conditional **AN**-stability and dynamical behaviour holds in the higher dimensional situation, since the nonautonomous stability criterion is based on the scalar test equation (5.32).

Chapter 7

Concluding Remarks

7.1 Summary of Thesis

The main objective of the thesis was to study the dynamical behaviour of linearized one-point collocation methods, used with constant stepsize, to discretize four classes of ordinary differential equations. The idea is to characterize the differences, if any, between some of the dynamical features of the discrete system and its continuous counterpart.

7.1.1 Scalar Parameter-dependent ODEs

Discretizing ODEs of the form

$$x' = f(x, \mu), \qquad x(0) = x_0 \tag{7.1}$$

where $\mu \in \mathbb{R}$ and $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, the thesis considered the question of the possible existence of spurious codimension-1 bifurcations in the methods.

Normal forms theory was used in the analysis.

It was established that no spurious saddle-node, transcritical, and pitchfork bifurcations can occur in linearized one-point collocation methods. In other words, any such bifurcation occurring in the methods must have resulted from a corresponding bifurcation in the originating ODE.

The period doubling bifurcation in discrete dynamical systems is a phenomenon that has no counterpart in continuous systems. It was discovered that spurious period doubling bifurcations can occur in all the methods with the exception of the linearized implicit midpoint method ($c_1 = 1/2$).

7.1.2 Planar Parameter-dependent ODEs

An ODE of the form (7.1) with $\mu \in \mathbb{R}$ and $f : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ was considered. The ODE was assumed to have an equilibrium point that undergoes a Hopf bifurcation at a known parameter value.

It was determined that, when using a linearized one-point collocation method to solve such an ODE, the corresponding fixed point in the resulting discrete system undergoes a Neimark-Sacker bifurcation (which is the discrete analog of the Hopf bifurcation) but, unless $c_1 = 1/2$, the bifurcation occurs at a different parameter value. The parameter value at which the bifurcation was found to be dependent on c_1 and the stepsize h.

The unforced and forced Van der Pol's oscillators were used to illustrate these results. In the case of the forced oscillator, the method of averaging was used to reduce the system to an autonomous one. The numerical results were consistent with theoretical expectations.

7.1.3 Planar Non-parameter-dependent ODEs

Concentrating on the autonomous ODE

$$x' = f(x), \qquad x(0) = x_0 \tag{7.2}$$

where $f : \mathbb{R} \to \mathbb{R}$, the issue of global asymptotic behaviour was considered.

As stated in Chapter 1, linearized one-point collocation methods do not generally admit spurious fixed points. However, the presence of a singularity in those methods for which $c_1 > 0$ may introduce spurious pole-type behaviour, and the numerical basin of attraction for the fixed points may differ from the true basin.

It was proved that if $f(x) = P(x)$, where $P(x)$ is a polynomial of degree at least 2, spurious pole-type behaviour will occur in the linearized one-point collocation method with $c_1 = 1/n$.

This is illustrated in $\mathbb{R}^2$ using a predator-prey model, in which the linearized implicit midpoint method depicts spurious pole-type behaviour.

7.1.4 Periodic Solutions - Nonautonomous ODEs

A scalar nonautonomous ODE of the form

$$x' = f(t, x), \qquad x(0) = x_0 \tag{7.3}$$

where $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and T-periodic in t, was considered. For the cases in which (7.3) is known to have a unique, asymptotically stable T-periodic solution, the main interest was to establish conditions under which the chosen methods exhibited the same dynamics.

The technique used is analogous to the Poincaré map, and is called stroboscopic sampling. Two cases were considered; $f = a(t)x + b(t)$ (linear case) and $f = a(t)g(x) + b(t)$ (nonlinear case), where $a(t)$ and $b(t)$ are C^1 and T-periodic and $g(x)$ is C^1 and generally nonlinear. The trivial cases, $a(t) = 0$ and $a(t) = -1$ in the linear category, were studied first.

Having established conditions for the existence of unique, asymptotically stable periodic solutions in the methods, it was investigated whether the conditions were linked to linear and nonlinear stability theory. It was proved that such a relationship exists; if a method is conditionally **AN**- or **BN**-stable, it will have a unique periodic solution that is asymptotically stable.

The study was taken to higher dimensions. Let

$$x' = A(t)x + b(t) \qquad x(0) = x_0 \tag{7.4}$$

where $A(t) \in \mathbb{R}^{m,m}$ is C^1 and T-periodic and $b(t) \in \mathbb{R}^m$ is also C^1 and T-periodic. The analysis was similar to the scalar case, and some of the difficulties that arise were highlighted.

7.2 Suggestions for Future Work

This thesis has raised a number of issues and questions that are worth investigating. The most natural direction in which to take this work is in generalizing the problems and methods used. The setting in which the thesis addresses some fundamental questions in the dynamics of numerics has been a simple one, yet it has enabled identification of areas to which further investigation could be directed.

Periodically forced systems play a very important role in practice. In the thesis, we considered the forced van der Pol's oscillator as an example and investigated the Neimark-Sacker bifurcation in linearized one-point collocation methods, used to discretize the averaged equations of this system. We would like to extend the investigation to consider the complete bifurcation analysis. For example, we would like to consider the effect of the proximity, in some places, of the saddle-node and Hopf bifurcations on the dynamics of numerics. Since, as it has been established in Chapter 4, the Hopf bifurcation does not occur at the same parameter value as it does in the differential equation, it would be interesting to study the consequences of this in the presence of a neighbouring bifurcation.

The link between the dynamics of the methods and the stability theory discussed in the thesis could naturally be investigated with other standard numerical methods such as Runge-Kutta and linear multistep methods. This also goes for the local bifurcations. Linearization of a general s-stage implicit Runge-Kutta method results in an s-stage *Rosenbrock method* (see Hairer & Wanner [15]), which, for the autonomous case, is given by

$$\begin{aligned} k_i &= f\left(X_n + \sum_{j=1}^{i-1} \alpha_{ij} k_j\right) + J \sum_{j=1}^{i} \gamma_{ij} k_j, \qquad i = 1, 2, \dots, s \\ X_{n+1} &= X_n + h \sum_{j=1}^{s} b_j k_j \end{aligned} \tag{7.5}$$

where $J = f'(x_0)$ and α_{ij}, γ_{ij}, b_i are the determining coefficients.

Applied to nonautonomous problems, these methods can be written

$$\begin{aligned} k_i &= f\left(t_n + \alpha_i h, X_n + \sum_{j=1}^{i-1} \alpha_{ij} k_j\right) + \gamma_i h \frac{\partial f}{\partial t}(t_n, X_n) + \frac{\partial f}{\partial x}(t_n, X_n) \sum_{j=1}^{i} \gamma_{ij} k_j \\ X_{n+1} &= X_n + h \sum_{j=1}^{s} b_j k_j, \end{aligned} \tag{7.6}$$

where $\alpha_i = \sum_{j=1}^{i-1} \alpha_{ij}$ and $\gamma_i = \sum_{j=1}^{i} \gamma_{ij}$.

Some Rosenbrock methods (7.5) or (7.6) can also be categorized as linearized collocation methods. This is the case if, in the derivation of the originating implicit Runge-Kutta method, a set of collocation points $\{t_n + c_i h\}$ $(i = 1, 2, \ldots, s)$ is chosen and the solution from t_n to t_{n+1} is approximated by a polynomial of degree s which satisfies the differential equation at the collocation points. Future work will attempt to base the dynamical approach study on such methods.

Also, it is known that explicit methods generally perform poorly, while Rosenbrock methods perform well, when used to solve stiff differential equations (Hairer & Wanner [15]). Future work will also attempt to apply linearized collocation methods to stiff equations, the aim being to ascertain if these methods can be used to solve stiff equations in general.

Finally, there has been increasing interest recently on local error control. It has been proved that, in most cases, spurious solutions cease to exist when local error control is used (see Aves *et al.* [3] and Sanz-Serna [26]). Furthermore, most modern numerical algorithms have built-in local error control. It would, therefore, be of great value to consider the problems and methods in this thesis within the variable time-stepping context. An understanding of the dynamics of fixed time-stepping methods is, however, necessary before embarking on variable time-stepping (see Stewart [29]).

Bibliography

[1] Ablowitz, M. J. & Segur, H. 1981 *Solitons and the Inverse Scattering Transform*. Philadelphia: SIAM.

[2] Arnold, V.I. 1983 *Geometric Methods in the Theory of Ordinary Differential Equations*. New York: Springer-Verlag.

[3] Aves, M.A., Griffiths, D.F. & Higham, D.J. 1997 Does error control suppress spuriosity? *SIAM J. Numer. Anal.* **34**(2), 756–778.

[4] Broomshead, D.S. & Iserles, A. (eds.) 1992 *The Dynamics of Numerics and Numerics of Dynamics*. Oxford: Clarendon Press.

[5] Brunner, H. 1997 The use of splines in the numerical solution of differential and Volterra integral equations. To appear: *Proc. of the CRM Workshops on Spline Functions and Theory of Wavelets*.

[6] Devaney, R. L. 1986 *Introduction to Chaotic Dynamical Systems.* Menlo Park: Benjamin-Cummings.

[7] Foster, E.A.D. & Khumalo, M. 1997 Transformation of local bifurcations under collocation methods. (*In preparation*)

[8] Foster, E.A.D. 1996 Boundary crises arising from saddle-node bifurcations in periodically forced models. *Int. J. Bifurcation and Chaos.* **6**(4), 647–659.

[9] Grammaticos, B. & Ramani, A. 1993 Investigating the integrability of discrete systems. *Int. J. Mod. Phy.* **B7**(21,22), 3551–3565.

[10] Griffiths, D.F., Sweby, P.K., & Yee, H.C. 1992 On spurious asymptotic solutions of explicit Runge-Kutta Methods. *IMA J. Numer. Anal.* **12**, 319–338.

[11] Guckenheimer, J. & Holmes, P. 1983 *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* New York: Springer-Verlag.

[12] Guillou, A. et Soulé, J.L. 1969 La résolution numérique des problémes différentiels aux conditions initiales par des méthodes de collocation. *R.I.R.I.* **3**, 17–44.

[13] Hairer, E., Iserles, A. & Sanz-Serna, J.M. 1989 Equilibria of Runge-Kutta Methods. *Numer. Math.* **58**, 243–254.

[14] Hairer, E., Norsett, S.P. & Wanner, G. 1993 *Solving Ordinary Differential Equations I: Non-stiff Problems* Second Edition. New York: Springer-Verlag.

[15] Hairer, E. & Wanner, G. 1991 *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems.* New York: Springer-Verlag.

[16] Hale, J., & Koçac, H. 1991 *Dynamics and Bifurcations.* New York: Springer-Verlag.

[17] Hartman, P. 1964 *Ordinary Differential Equations.* Baltimore: John Wiley & Sons, Inc.

[18] Humphries, A.R. 1993 Spurious solutions of numerical methods for initial value problems. *IMA J. Numer. Anal.* **13**, 263–290.

[19] Iserles, A. 1990 Stability and dynamics of numerical methods for nonlinear ordinary differential equations. *IMA J. Numer. Anal.* **10**, 1–30.

[20] Iserles, A., Peplov, A.T., & Stuart, A.M. 1991 A unified approach to spurious solutions introduced by time discretization. 1. Basic Theory. *SIAM J. Numer. Anal.* **28**(6), 1723–1751.

[21] Joshi, N. 1994 Singularity analysis and integrability for discrete dynamical systems. *J. Math. Anal. Appl.* **184**, 573–584.

[22] Khumalo, M. 1997 Dynamics of numerics of scalar nonautonomous equations with periodic solutions. *Submitted to: IMA J. Numer. Anal.*

[23] Lambert, J.D. 1991 *Numerical Methods for Ordinary Differential Systems.* New York: John Wiley & Sons.

[24] Loscalzo, F.R. & Talbot, T.D. 1967 Spline function approximations for solutions of ordinary differential equations. *SIAM J. Numer. Anal.* **4**, 433–445.

[25] Pomeau, Y. & Manneville, P. 1980 Intermittent transition to turbulence in dissipative dynamical systems. *Commun. Math. Phys.* **74**, 189–197.

[26] Sanz-Serna, J.M. 1992 Numerical ordinary differential equations versus dynamical systems. In Broomshead, D.S. & Iserles, A. (ed.) *The Dynamics of Numerics and the Numerics of Dynamics.* Oxford: Clarendon Press.

[27] Sanchez, D. 1968 *Ordinary Differential Equations and Stability Theory: An Introduction.* San Francisco: W.H. Freeman and Co.

[28] Stein, O. 1997 Bifurcations of hyperbolic fixed points for explicit Runge-Kutta methods. *IMA J. Numer. Anal.* **17**, 151–175.

[29] Stewart, I. 1992 Warning — handle with care! *Nature* **355**, 16–17.

[30] Stuart, A. M. 1989 Linear instability implies spurious periodic solutions. *IMA J. Numer. Anal.* **9**, 465–488.

[31] Stuart, A.M. & Humphries, A.R. 1996 *Dynamical Systems and Numerical Analysis.* New York: Cambridge University Press.

[32] Wiggins, S. 1990 *Introduction to Applied Nonlinear Dynamical Systems and Chaos.* New York: Springer-Verlag.

[33] Yee, H. C., & Sweby, P.K. 1994 Global asymptotic behavior of iterative implicit schemes. *Int. J. Bifurcation and Chaos.* **4**(6), 1579–1611.

[34] Yee, H.C., & Sweby, P.K. 1997 Dynamics of numerics and spurious behaviors in CFD computations. *RIACS Technical Report 97.06.*

[35] Yee, H.C., Sweby, P.K., & Griffiths, D.F. 1991 Dynamical approach study of spurious steady-state numerical solutions of nonlinear differen-

tial equations.I.The dynamics of time discretization and its implications for algorithm development in CFD. *J. Comp. Phys.* **97**, 249–310.

[36] Yee, H.C. & Sweby, P.K. 1992 Dynamical approach study of spurious steady-state numerical solutions of nonlinear differential equations: II. *RNR Technical Report RNR-92-008.*