

HIGH ORDER COMPACT SCHEME FOR  
DISCONTINUOUS DIFFERENTIAL  
EQUATIONS

by

JAMES LOUIS THOMPSON

Presented to the Faculty of the Graduate School of  
The University of Texas at Arlington in Partial Fulfillment  
of the Requirements  
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

MAY 2011

Copyright © by JAMES LOUIS THOMPSON 2011

All Rights Reserved

## ACKNOWLEDGEMENTS

I would like to express sincere appreciation to Dr, Chaoqun Liu for his guidance through the process of preparing for and writing of this dissertation. Dr. Liu's knowledge and patience are extremely helpful. I would also like to thank the committee members for contributing their time reviewing this dissertation.

I want to thank my parents for always encouraging me in academic matters, and bearing much of the financial cost.

Finally, I would to thank my wife and daughter for tolerating my absence from them during the preparation of this dissertation

April 18, 2011

ABSTRACT

HIGH ORDER COMPACT SCHEME FOR  
DISCONTINUOUS DIFFERENTIAL  
EQUATIONS

JAMES LOUIS THOMPSON, PhD

The University of Texas at Arlington, 2011

Supervising Professor: CHAOQUN LIU

A high order implicit second derivative compact method is given which is similar the Adams-Moulton method, but requiring only two steps for sixth order. This method is used in both predictor-corrector and Newton's method formulations, and although the compact scheme is not A-stable or stiffly stable, it's region of stability is over six times greater than the Sixth order Adams-Moulton method. This compact method has a small truncation error coefficient, and is more accurate than Enright's method within the region of stability.

Fourth and sixth order explicit compact methods are derived as well, and the three step sixth order explicit method is used as the predictor for the harmonic oscillator equation with Coulomb damping. Consistency and rate of convergence conditions are derived for these compact methods, and convergence is proved as well. The region of stability is plotted for the sixth order implicit case.

The two step sixth order implicit compact method is compared against the five step stiffly-stable Enright method and the five step Adams-Moulton scheme on three test problems, and is

shown to be more accurate than Enright's method, and has better accuracy and is more stable than Adams-Moulton.

The predictor-corrector compact formulation is tested on two Coulomb friction problems, and the difficulty caused by the discontinuous right-hand side is avoided by breaking the problem into segments between the discontinuities.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	iii
ABSTRACT.....	iv
LIST OF ILLUSTRATIONS.....	viii
LIST OF TABLES.....	ix
Chapter	Page
1. FRICTIONAL DAMPING.....	1
1.1 Viscous Friction .....	1
1.2 Coulomb Friction .....	3
1.3 Forced Vibration.....	6
1.4 Numerical Approach.....	9
2. THE COMPACT METHOD .....	13
2.1 Hermite Polynomial Interpolation .....	13
2.2 Explicit Compact Formulas.....	16
2.3 Implicit Compact Method.....	19
2.4 Truncation Error.....	20
3. CONSISTENCY, CONVERGENCE, and STABILITY .....	23
3.1 Consistency and Speed of Conversion .....	23
3.2 Convergence.....	28
3.3 Stability.....	31
4. NUMERICAL EXAMPLES.....	35
4.1 Comparing Compact, Enright, and Adams-Moulton Methods.....	35

4.2 The Compact Method for Problems with Coulomb Friction.....	40
4.3 Conclusions.....	43
REFERENCES.....	45
BIOGRAPHICAL INFORMATION.....	47

## LIST OF ILLUSTRATIONS

Figure	Page
1.1 Error in Runge-Kutta calculation for a simple Coulomb damping problem.....	10
3.1 Region of stability for the two step implicit compact method .....	34
4.1 Absolute value of the error for the two step compact implicit method applied to problem (4.5).....	37
4.2 Absolute value of the error for the four step Enright method for to problem (4.5).....	37
4.3 Absolute value of the error for the two step implicit compact method for problem (4.7).....	39
4.4 Absolute value of the error for Enright's method applied to problem (4.7).....	39
4.5 Compact method solutions for position and velocity for the unforced Coulomb damping problem.....	41
4.6 Positional error for the unforced Coulomb friction problem .....	42
4.7 Velocity error for the unforced Coulomb friction problem.....	42
4.8 Solution for position in problem (4.12) with viscous and Coulomb damping.....	43
4.9 Solution for velocity in problem (4.12) with viscous and Coulomb damping.....	43



## LIST OF TABLES

Table	Page
4.1 Comparison of the absolute value of error for sixth order implicit compact, Enright, and Adams-Moulton methods for problem (4.3).....	36
4.2 Comparison of the absolute value of error for sixth order implicit compact, Enright, and Adams-Moulton methods for problem (4.5).....	36

CHAPTER 1  
FRICTIONAL DAMPING

This chapter discusses frictional damping in simple physical systems. The examples are all one dimensional, and the equations of motion are derived for both viscous and Coulomb friction. The purpose of this chapter is to provide motivational background for the differential equations studied in the following chapters, and is intended as a brief survey rather than an exhaustive study.

1.1 Viscous Friction

Viscosity may be described roughly as the shear forces existing in moving fluids (Feynman, 1963). There are many examples of viscous forces, including air resistance to slowly moving bodies, and pistons sliding in lubricated cylinders. Consider a plate sliding on a flat surface which has been coated with a liquid. It is known that the force on the plate required to maintain a constant velocity is (Feynman, 1963; Hutton, 1981)

$$F = \eta \frac{v}{d} A, \quad (1.1)$$

where  $\eta$  is the coefficient of viscosity,  $v$  is the velocity,  $d$  is the thickness of the fluid, and  $A$  is the surface area of the plate contacting the fluid.

If the free end of a securely anchored spring is attached to the plate, the spring/plate system displaced a distance  $x_0$  in the positive  $x$  direction, and the plate then released, how will the viscosity of the fluid affect the motion? Applying Newton's second law to this system gives

$$mx'' = -kx - cx', \quad (1.2)$$

where the constant  $c$  is defined as

$$c \equiv \eta \frac{A}{d}. \quad (1.3)$$

Equation (1.2) is usually written as

$$x'' + \frac{c}{m}x' + \frac{k}{m}x = 0. \quad (1.4)$$

The linear homogeneous differential equation (1.4) can be solved by the usual techniques.

Defining the undamped frequency  $\omega_0$  as

$$\omega_0 \equiv \sqrt{\frac{k}{m}}, \quad (1.5)$$

and the damping factor as (Hutton, 1981; Jacobsen and Ayre, 1958)

$$\zeta \equiv \frac{c}{2m\omega_0}, \quad (1.6)$$

the roots of the auxiliary equation for (1.4) are

$$s_{1,2} = \omega_0 \left( -\zeta \pm \sqrt{\zeta^2 - 1} \right). \quad (1.7)$$

The value of  $\zeta$  gives three distinct solutions to equation (1.4). These solutions are called overdamped, critically damped, and underdamped for  $\zeta > 1$ ,  $\zeta = 1$ , and  $\zeta < 1$  respectively.

$$\zeta > 1, \quad x(t) = a_1 e^{-\omega_0(\zeta - \sqrt{\zeta^2 - 1})t} + a_2 e^{-\omega_0(\zeta + \sqrt{\zeta^2 - 1})t} \quad (1.8)$$

$$\zeta = 1, \quad x(t) = (a_3 + a_4 t) e^{-\omega_0 t} \quad (1.9)$$

$$\zeta < 1, \quad x(t) = e^{-\zeta \omega_0 t} \left( a_5 \sin \sqrt{1 - \zeta^2} \omega_0 t + a_6 \cos \sqrt{1 - \zeta^2} \omega_0 t \right), \quad (1.10)$$

where the constants  $a_j$  are determined by the initial conditions.

Equations (1.8), (1.9), and (1.10) show the effects of viscous damping on the motion, namely that the oscillatory motion has a lower frequency than the undamped case, and the amplitude of the oscillation decays to zero exponentially. Additionally, if the damping is strong enough ( $\zeta \geq 1$ ) there is no oscillatory motion at all, and the plate approaches the static equilibrium point  $x = 0$  asymptotically.

## 1.2 Coulomb Friction

If the viscous fluid is removed, the plate must slide on the dry table surface. This type of frictional force is then called Coulomb friction, although it is sometimes called dry, solid, or constant friction. Coulomb frictional force always opposes the motion, but can never cause motion.

The force required to maintain constant plate motion is

$$F_f = \mu N . \quad (1.11)$$

Here  $N$  is the normal force between the table and the plate, and  $\mu$  is the coefficient of friction. Although not strictly true (Feynman, 1962), the coefficient of friction  $\mu$  is assumed to be independent of velocity, with value between zero and one. If the table surface is horizontal, and there is no vertical force other than gravity, then the normal force  $N$  equals the weight  $W$  of the plate.

If the plate velocity is zero, the frictional force is given by (1.11) only if the magnitude of the net applied horizontal force  $F_A$  is greater than  $\mu N$ . Otherwise the frictional force is

$$F_f = -F_A . \quad (1.12)$$

Replacing viscous damping in the spring/plate system of section (1.1) with Coulomb friction changes the equation of motion (1.4) to

$$x'' + \frac{k}{m}x - \frac{F_f}{m} = 0 , \quad (1.13)$$

where  $F_f$  is given by

$$F_f = -\mu W [\text{sgn}(x') - (1 - |\text{sgn}(x')|)\text{sgn}(x)] \quad (1.14)$$

for non-zero velocity, or for zero velocity with

$$|kx(t)| > \mu W . \quad (1.15)$$

The signing function in (1.14) is defined as

$$\begin{aligned}
\operatorname{sgn}(z) &= 1, & z > 0 \\
\operatorname{sgn}(z) &= 0, & z = 0 \\
\operatorname{sgn}(z) &= -1 & z < 0
\end{aligned} \tag{1.16}$$

If the velocity is zero and (1.15) is not satisfied,

$$F_f = kx(t). \tag{1.17}$$

The second term in (1.14) is often ignored in the literature (Jacobsen and Ayre, 1958), but it is necessary if (1.13) is to obey the second law at the turning points. If the initial conditions are

$$\begin{aligned}
x(0) &= a \\
x'(0) &= 0,
\end{aligned} \tag{1.18}$$

with  $a$  positive and

$$a > \frac{\mu W}{k}, \tag{1.19}$$

then differential equation (1.13) and initial conditions (1.18) comprise a non-linear initial value problem. The non-linearity comes from the jump discontinuity of the Coulomb frictional force at each point in time where the velocity  $x'(t_i)$  changes direction. Equation (1.14) shows that the frictional force instantaneously changes sign whenever the plate stops. The magnitude of the jump is  $2\mu W$  at each turning point,  $\mu W$  if the plate comes to rest at  $x = 0$ , and  $\mu W \pm |kx(t_j)|$  at the remaining points of the dead band (points where the spring force is too small to cause motion). The physics of the spring/plate and similar problems means that the set of all times  $t_i$  where the discontinuities of  $x''$  occur is a set of measure zero.

The initial value problem defined by (1.13) and (1.18) can be solved with conventional techniques in a piecewise manner, but a Laplace transform solution is given here instead (Pipes, 1970; Hutton, 1981). The differential equation (1.13) may be written as

$$x''(t) + \omega_0^2 x(t) = \mu gh(t). \tag{1.20}$$

Here  $\omega_0$  is defined by (1.5), and the coefficient of  $h(t)$  comes from the fact that

$$W = mg, \quad (1.21)$$

with  $g$  defined as the acceleration of gravity. The function  $h(t)$  is a square wave with period  $2T$  and amplitude one, and may also be expressed as a sum of Heaviside functions

$$h(t) = u(t) + 2 \sum_{n=1}^N (-1)^n u(t - nT). \quad (1.22)$$

The Laplace transforms of  $x(t)$  and  $x''(t)$  are

$$L\{x(t)\} \equiv X(s), \quad (1.23)$$

$$\begin{aligned} L\{x''(t)\} &\equiv s^2 X(s) - sx(t=0) - x'(t=0) \\ &= s^2 X(s) - sa. \end{aligned} \quad (1.24)$$

The shifting theorem

$$L\{f(t - \tau)u(t - \tau)\} = e^{-s\tau} F(s) \quad (1.25)$$

is used to transform  $h(t)$ ,

$$L\{h(t)\} = \frac{1}{s} + \frac{2}{s} \sum_{n=1}^N (-1)^n e^{-nTs}. \quad (1.26)$$

The transformed differential equation is

$$(s^2 + \omega_0^2)X(s) = sa + \frac{\mu g}{s} \left( 1 + 2 \sum_{n=1}^N (-1)^n e^{-nTs} \right). \quad (1.27)$$

The solution for the transformed displacement is

$$X(s) = \frac{sa}{s^2 + \omega_0^2} + \frac{\mu g}{s(s^2 + \omega_0^2)} + 2\mu g \sum_{n=1}^N (-1)^n \frac{e^{-nTs}}{s(s^2 + \omega_0^2)}. \quad (1.28)$$

The inverse transform  $x(t)$  of  $X(s)$  is

$$x(t) = a \cos \omega_0 t + \frac{\mu g}{\omega_0^2} (1 - \cos \omega_0 t) + \frac{2\mu g}{\omega_0^2} \sum_{n=1}^N (-1)^n [1 - \cos \omega_0 (t - nT)] u(t - nT) \quad (1.29)$$

This motion will continue until the velocity goes to zero inside the dead band defined by

$$-\frac{\mu W}{k} \leq x \leq \frac{\mu W}{k}, \quad (1.30)$$

after which the spring no longer exerts enough force to overcome friction. This assumes that the amplitude  $x(t)$  decreases with time. Consider a time  $t$  such that  $t \in [0, T]$ , and compare the amplitude with that of  $x(t)$  with  $t \in [T, 2T]$ . After the initial displacement  $a$ , the first maximum displacement during the first half cycle occurs at  $t = T$ , where

$$x(T) = -a + 2 \frac{\mu g}{\omega_0^2}. \quad (1.31)$$

The second half cycle displacement extreme occurs at  $t = 2T$ ,

$$x(2T) = a - 4 \frac{\mu g}{\omega_0^2}, \quad (1.32)$$

a decrease in amplitude of  $2 \frac{\mu g}{\omega_0^2}$ .

The above analysis of the spring/plate system with Coulomb friction shows that the amplitude of the motion decays linearly with time, and vibrates at the same frequency as a frictionless oscillator. The most important difference between Coulomb and viscous damping is that the jump discontinuities of the Coulomb friction term make the differential equation nonlinear.

### 1.3 Forced Vibrations

In sections 1.1 and 1.2 an external force was applied to the spring/mass system to give initial displacement to the system, but thereafter the damping forces were the only external forces present. Consider now the application of a sinusoidal forcing function will then have equation of motion

$$mx'' + cx' + kx = F_E \sin \omega t. \quad (1.33)$$

This is normally rewritten as

$$x'' + 2\omega_0\zeta x' + \omega_0^2 x = X_0\omega_0^2 \sin \omega t, \quad (1.34)$$

where

$$X_0 = \frac{F_E}{k}, \quad (1.35)$$

and the undamped frequency  $\omega_0$  and damping factor  $\zeta$  are defined as before (see equations (1.5) and (1.6) respectively). The inhomogeneous differential equation (1.34) is linear; hence the solution is the sum of the homogenous solution and a particular solution. The homogeneous solutions are given by (1.8), (1.9), and (1.10), and since these solutions all decay to zero, the only remaining part of the solution is the particular solution. The particular solution is therefore called the steady state solution. The steady state solution is (Hutton, 1981; Jacobsen and Ayre, 1958)

$$x(t) = \frac{X_0}{\sqrt{(1-r^2)^2 + (2\zeta r)^2}} \sin(\omega t - \phi), \quad (1.36)$$

where the phase angle  $\phi$  is defined as

$$\phi = \text{ArcTan}\left(\frac{2\zeta r}{1-r^2}\right). \quad (1.37)$$

The variable  $r$  is the forcing frequency  $\omega$  divided by  $\omega_0$ , and is called the frequency ratio. The maximum value of  $x(t)$  occurs when

$$r = \sqrt{1-2\zeta^2}, \quad (1.38)$$

as long as

$$\zeta \leq \frac{\sqrt{2}}{2}. \quad (1.39)$$

This maximum amplitude is

$$x_{\max} = \frac{X_0}{2\zeta\sqrt{1-\zeta^2}}. \quad (1.40)$$



Equation (1.40) has the expected pole at resonance ( $r = 1$ ) if  $\zeta = 0$ , but for non-zero values of  $\zeta$ , the amplitude is a bounded function of  $r$ . Thus viscous damping removes the resonant pole of the undamped oscillator, and shifts the maximum response to values of  $r$  between zero and one.

The equation of motion for Coulomb damping with a forcing function is

$$mx'' + kx - F_f = F_E \sin \omega t, \quad (1.41)$$

with  $F_f$  having the same definition as in the unforced Coulomb damping case. The normal form for equation (1.41) is

$$x'' + \omega_0^2(x - X_f) = X_0 \omega_0^2 \sin \omega t, \quad (1.42)$$

where  $\omega_0$  and  $X_0$  are defined as in the viscous case, and  $X_f$  is defined by

$$X_f \equiv \frac{F_f}{k}. \quad (1.43)$$

The solution to (1.42) can have many forms (Den Hartog, 1985). There can be steady state solutions without stops in motion, or with one or more stops per half cycle. Den Hartog derived exact solutions to (1.42) for the cases of zero and one stops in each half cycle. Den Hartog's results will not be reproduced here, but the results will be summarized.

First, if a steady state solution exists, it will have the same frequency as the forcing function. Second, the motion may stop one or more times per half cycle. The last, and most surprising aspect of the solution is that for values of  $F_f$  and  $F_E$  such that

$$\frac{F_f}{F_E} \leq \frac{\pi}{4}, \quad (1.44)$$

the solution is unbounded at resonance. This result can be understood by looking at the difference in work done by the forcing function and the work done by friction. Since both are proportional to the amplitude, if  $F_f$  is less than some fraction of  $F_E$ , the energy input is greater

than the energy dissipated regardless of the amplitude, therefore the amplitude grows without bound. For viscous damping the dissipation is proportional to amplitude squared, and therefore as the amplitude grows the dissipation quickly catches up to the energy input, and equilibrium is achieved.

#### 1.4 Numerical Approach

The non-linear nature of Coulomb friction problems makes numerical solutions problematic (Gear, 1971). This can be illustrated by the simple unforced Coulomb problem (1.13), with  $\omega_0 = 1$ , and  $\mu g = .2$  (Thompson and Liu, 2004). The exact solution to this problem is

$$\begin{aligned} x(t) &= .8\cos(t) + .2 & 0 \leq t < \pi \\ x(t) &= .4\cos(t) - .2 & \pi \leq t < 2\pi \\ x(t) &= .2 & t \geq 2\pi, \end{aligned} \tag{1.45}$$

$$\begin{aligned} x'(t) &= -.8\sin(t) & 0 \leq t < \pi \\ x'(t) &= -.4\sin(t) & \pi \leq t < 2\pi \\ x'(t) &= 0 & t \geq 2\pi. \end{aligned} \tag{1.46}$$

Apply the fourth order Runge-Kutta explicit method in the usual way, including the correct Coulomb friction model (1.14), (1.15), and (1.16), to this differential equation. Figure 1.1 shows the normalized error in the numerical solution as a function of time for  $t \in [0,20]$ . Figure 1.1 clearly shows the rapid increase in error after the first discontinuity. The error is normal during  $0 \leq t < \pi$  (approximately  $10^{-6}$ ), but jumps three orders of magnitude after the first discontinuity. The error increases asymptotically after the second turning point; in fact the numerical solution is decaying towards zero for  $t \geq 2\pi$ . Any numerical method used for problems with Coulomb friction must overcome the effects of these discontinuities. These difficulties suggest breaking the domain into discrete intervals, with each interval bounded by the time at which discontinuity occurs (Gear, 1981). Although the differential equations are well behaved in the interior of each

interval, this approach will require automatic detection of each discontinuity. For Coulomb friction problems, this can be done by looking for sign changes in velocity at each point of the calculation.

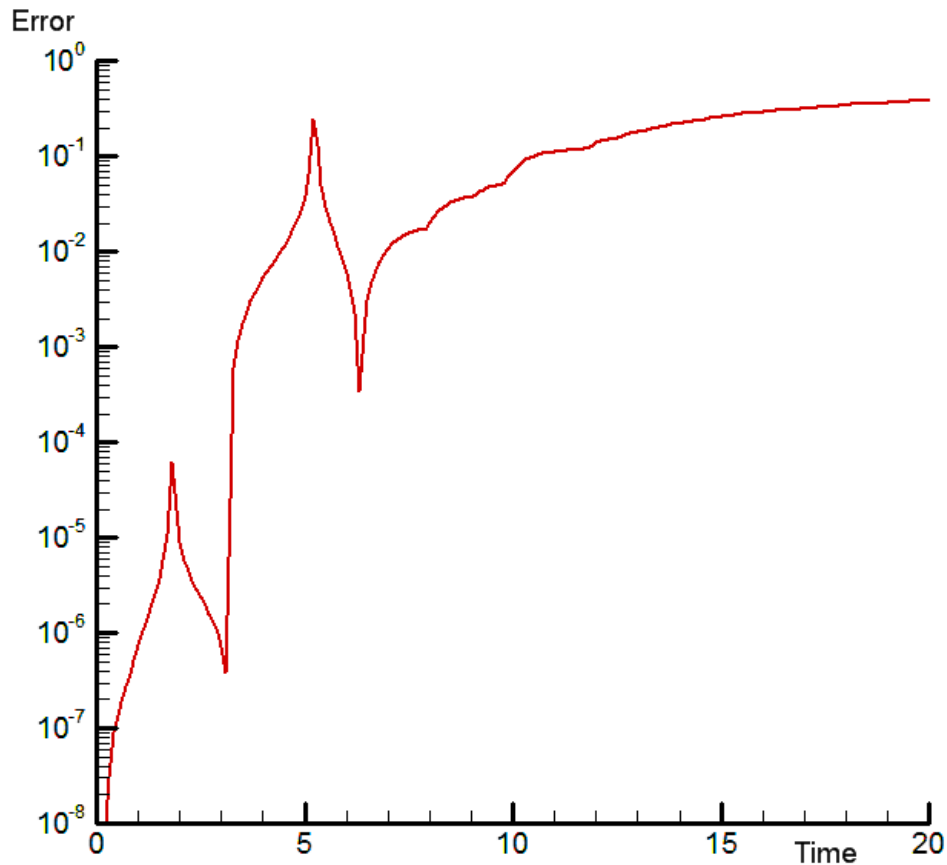


Figure 1.1 Error in Runge-Kutta calculation for a simple Coulomb damping problem. The computation was done in double precision with  $h = 0.1$

The code must then go back to the previous grid point and find the time where the velocity is zero; a time which is usually between grid points. This time will then begin the calculation as a new problem, with initial conditions given by

$$\begin{aligned} x(t_\alpha) &\equiv x_\alpha \\ x'(t_\alpha) &\equiv 0. \end{aligned} \tag{1.47}$$

The accuracy of the zero velocity time will affect the error in the following interval, and therefore a high order numerical method, a grid refinement, or a combination of both is necessary.

Simple single step methods such as the explicit fourth order Runge-Kutta or Obrechhoff's A-stable second derivative method (Obrechhoff, 1942) could easily be implemented for these problems. More complex methods involving blocked Runge-Kutta schemes give high order and good stability (Shampine, 2011; Verner, 2008). Shampine's explicit Runge-Kutta pair uses eight equally spaced points in the span of each step, and are eighth order at interior points and ninth order at the step grid points. The average radius of the stability region for Shampine's scheme is approximately 4.3. The blocked Runge-Kutta methods would require modification to detect the points of discontinuity in Coulomb friction problems, and more generally, matrix methods pose additional difficulties during the detection process. Hence the goal in this work is higher order compact methods (Lele, 1992). A two-step, sixth order implicit compact method (Thompson and Liu, 2004) will give good accuracy and stability with less complexity than block Runge-Kutta methods, and is the method used in this work. This method can be used in a predictor-corrector scheme with a sixth order explicit compact corrector, or combined with Newton's method in a more straight forward approach. Although the compact schemes are not A-stable, the stability region of the implicit two step method is considerably larger than either the fourth order explicit Runge-Kutta method, sixth order implicit Adams-Moulton method, and the ninth order Shampine scheme.

In cases where the spring constant is large enough to require stiffly stable methods, Enright has published methods of the form (Enright, 1974)

$$y_{n+1} = y_n + h \sum_{j=0}^k b_j y'_{n+1-j} + h^2 \gamma_0 y''_{n+1}, \quad (1.48)$$

where  $k$  is the number of steps. These formulae are of order  $k + 2$ , and are A-stable up through two steps, and stiffly-stable for  $3 \leq k \leq 7$ . Thus a sixth order Enright scheme would be a four-step method, and is later shown to be less accurate than the sixth order two-step compact scheme.

The compact method is derived in the same way as the formulae for the Adams methods, except Hermite polynomials are used to interpolate  $f(t, y(t))$  rather than the Lagrange interpolation used in the Adams schemes. The function  $f(t, y(t))$  is the right hand side of

$$y'(t) = f(t, y(t)). \quad (1.49)$$

Hermite polynomials have the property

$$\begin{aligned} H(t_i) &= f(t_i, y(t_i)) \\ H'(t_i) &= f'(t_i, y(t_i)), \end{aligned} \quad (1.50)$$

thereby doubling the order of the polynomial. This gives doubles the order of the Adams methods for a given number of steps.

The derivation of both explicit and implicit compact equations is done in Chapter 2. Chapter 3 looks at consistency, convergence, and stability for the compact method, and finally in Chapter 4 the compact method is compared to Adams-Moulton and Enright methods, and then applied to differential equations with Coulomb and viscous friction.

## CHAPTER 2

### THE COMPACT METHOD

In this chapter both two and three step explicit and two step implicit compact schemes are derived for first order differential equations of the form

$$y'(t) = f(t, y(t)), \quad y(a) = y_0, \quad (2.1)$$

where the independent variable  $t \in [a, b]$ .

The function  $f(t, y(t))$  is interpolated with Hermite polynomials, and these polynomials are used to approximate  $f(t, y(t))$  in

$$\begin{aligned} y(t_{i+1}) - y(t_i) &= \int_{t_i}^{t_{i+1}} y'(t) dt \\ &= \int_{t_i}^{t_{i+1}} f(t, y(t)) dt. \end{aligned} \quad (2.2)$$

The first section discusses the application of Hermite polynomials to the derivation of the compact schemes, the second section contains the derivation of the difference equations for both two and three step explicit schemes, and section three in this chapter is the derivation of the difference equation for the two step implicit compact scheme.

#### 2.1 Hermite Polynomial Interpolation

Since it may not be possible to perform the integration in equation (2.2), it may be necessary to seek approximate solutions. One possible approach is the polynomial interpolation of  $f(t, y(t))$  on some interval of time. This interpolation scheme is the essence of well known multistep methods such as Adams-Bashforth and Adams-Moulton. In the Adams methods the coefficients of each power of the independent variable in

$$P(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_k t^k \quad (2.3)$$

are chosen such that

$$P(t_i) = f(t_i, y(t_i)) \quad (2.4)$$

whenever  $t_i \in \{t_{n+1}, t_n, t_{n-1}, \dots, t_{n+1-m}\}$  for the implicit Adams-Moulton scheme. For the explicit Adams-Bashforth case  $t_i \in \{t_n, t_{n-1}, \dots, t_{n+1-m}\}$ . In Both cases the index  $m$  denotes the number of prior grid steps whose information is incorporated into the solution at  $t_{n+1}$ . It will be shown in the work that follows that adding the requirement

$$P'(t_i) = f'(t_i, y(t_i)) \quad (2.5)$$

reduces the error due to approximation of  $f(t, y(t))$  over that where only (2.4) is used. The addition of (2.5) to the interpolation of  $f(t, y(t))$  is uniquely satisfied by the Hermite polynomials

$$f(x) = H_{2n+1}(x) + \frac{(x-x_0)^2 \Lambda (x-x_n)^2}{(2n+2)!} f^{(2n+2)}(\zeta), \quad (2.6)$$

for some  $\zeta(x) \in (a, b)$  (Burden and Faires, 2005).

The Hermite polynomials of equation (2.6) are now used to construct numerical schemes for the solution of the first order differential equation (2.1). The interpolation will be done on a subset of the uniform grid defined by

$$t_n = a + nh, \quad n \in \{0, 1, K, N\}, \quad (2.7)$$

with

$$h = \frac{b-a}{N}. \quad (2.8)$$

The numerical solution  $y_{n+1}$  at the grid point  $t_{n+1}$  will be calculated by interpolating  $f(t, y(t))$  on the preceding  $m$  grid points for explicit methods, and on the grid points  $t_{n+1-m}, K, t_{n-1}, t_n, t_{n+1}$  for implicit methods. The Hermite polynomial for the  $m$ -step explicit method is

$$H_{2m-1}(t) = a_0 + a_1 t + a_2 t^2 + \Lambda + a_{2m-1} t^{2m-1}. \quad (2.9)$$

The  $2m$  coefficients  $a_0, a_1, K, a_{2m-1}$  can be obtained by solution of the equations

$$a_0 + a_1 t_n + a_2 t_n^2 + \Lambda + a_{2m-1} t_n^{2m-1} = f_n$$

$$a_1 + 2a_2 t_n + \Lambda + (2m-1)a_{2m-1} t_n^{2m-2} = f'_n$$

$$a_0 + a_1 t_{n-1} + a_2 t_{n-1}^2 + \Lambda + a_{2m-1} t_{n-1}^{2m-1} = f_{n-1}$$

$$a_1 + 2a_2 t_{n-1} + \Lambda + (2m-1)a_{2m-1} t_{n-1}^{2m-2} = f'_{n-1}$$

\(\mathfrak{N}\)

$$a_0 + a_1 t_{n+1-m} + a_2 t_{n+1-m}^2 + \Lambda + a_{2m-1} t_{n+1-m}^{2m-1} = f_{n+1-m}$$

$$a_1 + 2a_2 t_{n+1-m} + 3a_3 t_{n+1-m}^2 + \Lambda + (2m-1)a_{2m-1} t_{n+1-m}^{2m-2} = f'_{n+1-m}. \quad (2.10)$$

The notation  $f_k$  in (2.10) is an abbreviation for  $f(t_k, y(t_k))$ .

The interpolating polynomial for implicit  $m$ -step schemes is given by

$$H_{2m+1}(t) = a_0 + a_1 t + a_2 t^2 + \Lambda + a_{2m+1} t^{2m+1}. \quad (2.11)$$

For a given step size, the order of polynomial (2.11) is two greater than the order of polynomial (2.10) used for explicit methods. For implicit methods, the  $2m+2$  coefficients  $a_0, a_1, \Lambda, a_{2m+1}$  are determined from

$$a_0 + a_1 t_{n+1} + a_2 t_{n+1}^2 + \Lambda + a_{2m+1} t_{n+1}^{2m+1} = f_{n+1}$$

$$a_1 + 2a_2 t_{n+1} + \Lambda + (2m+1)a_{2m+1} t_{n+1}^{2m} = f'_{n+1}$$

$$a_0 + a_1 t_n + a_2 t_n^2 + \Lambda + a_{2m+1} t_n^{2m+1} = f_n$$

$$a_1 + 2a_2 t_n + \Lambda + (2m+1)a_{2m+1} t_n^{2m} = f'_n$$

\(\mathfrak{N}\)

$$a_0 + a_1 t_{n+1-m} + a_2 t_{n+1-m}^2 + \Lambda + a_{2m+1} t_{n+1-m}^{2m+1} = f_{n+1-m}$$

$$a_1 + 2a_2 t_{n+1-m} + 3a_3 t_{n+1-m}^2 + \Lambda + (2m+1)a_{2m+1} t_{n+1-m}^{2m} = f'_{n+1-m}. \quad (2.12)$$



The solutions for the coefficients  $a_k$  in (2.10) and (2.12) are used to derive the compact methods.

Using equation (2.9) in the right hand side of (2.2) generates the explicit compact methods,

$$y(t_{i+1}) = y(t_i) + \int_{t_i}^{t_{i+1}} H_{2m-1}(t) dt, \quad (2.13)$$

and the implicit schemes are derived by using (2.11),

$$y(t_{i+1}) = y(t_i) + \int_{t_i}^{t_{i+1}} H_{2m+1}(t) dt. \quad (2.14)$$

The Adams schemes are derived using Newton's backward-difference formulas rather than finding the polynomial coefficients directly (Atkinson, 1989; Burden and Faires, 2005). While the Hermite polynomials have a Newton divided difference interpolation formula, this approach is not as advantageous as in the Adams methods and is not used here.

## 2.2 Explicit Compact Formulas

The formulas for explicit two and three step methods are derived in this section, with rational coefficient expressions in both cases.

First, equation (2.10) is simplified by shifting the time axis such that the subscript  $n$  corresponds to  $t = 0$ . This transformation reduces equation (2.10) to

$$a_0 = f_n$$

$$a_1 = f'_n$$

$$a_2 h^2 + \Lambda + a_{2m-1} (-h)^{2m-1} = f_{n-1} + h f'_n - f_n$$

$$-2a_2 h + \Lambda + (2m-1)a_{2m-1} (-h)^{2m-2} = f'_{n-1} - f'_n$$

$\Lambda$

$$a_2 (m-1)^2 h^2 - a_3 (m-1)^3 h^3 + \Lambda + a_{2m-1} (-(m-1)h)^{2m-1} = f_{n+1-m} - f_n + (m-1)h f'_n$$

$$-2a_2 (m-1)h + 3a_3 (m-1)^2 h^2 + \Lambda + (2m-1)a_{2m-1} (-(m-1)h)^{2m-2} = f'_{n+1-m} - f'_n \quad (2.15)$$

After solving (2.15) for the coefficients  $a_0, a_1, \dots, a_{2m-1}$ , the explicit compact method is determined from (2.13)

$$y(t_{i+1}) = y(t_i) + \int_0^h (a_0 + a_1 t + \Lambda + a_{2m-1} t^{2m-1}) dt,$$

$$y_{i+1} = y_i + h a_0 + \frac{h^2}{2} a_1 + \frac{h^3}{3} a_2 + \Lambda + \frac{h^{2m}}{2m} a_{2m-1}. \quad (2.16)$$

This gives the numerical scheme as a function of  $f$  evaluated at the previous  $m$  grid points.

For a two step method, equation (2.15) becomes

$$a_0 = f_n$$

$$a_1 = f'_n$$

$$h^2 a_2 - h^3 a_3 = f_{n-1} + h f'_n - f_n$$

$$-2h a_2 + 3h^2 a_3 = f'_{n-1} - f'_n. \quad (2.17)$$

The solution for  $a_2$  and  $a_3$  is

$$a_2 = \frac{1}{h^2} \{-3f_n + 3f_{n-1} + 2hf'_n + hf'_{n-1}\},$$

$$a_3 = \frac{1}{h^3} \{-2f_n + 2f_{n-1} + hf'_n + hf'_{n-1}\}. \quad (2.18)$$

Inserting  $a_0, a_1, a_2$ , and  $a_3$  into (2.16) gives the two step explicit compact method,

$$y_{n+1} = y_n + \frac{h}{12} [-6f_n + 18f_{n-1} + h(17f'_n + 7f'_{n-1})]. \quad (2.19)$$

The coefficient equations for the three step explicit scheme is

$$a_0 = f_n$$

$$a_1 = f'_n$$

$$h^2 a_2 - h^3 a_3 + h^4 a_4 - h^5 a_5 = -f_n + f_{n-1} + hf'_n$$

$$\begin{aligned}
-2ha_2 + 3h^2a_3 - 4h^3a_4 + 5h^4a_5 &= -f'_n + f'_{n-1} \\
4h^2a_2 - 8h^3a_3 + 16h^4a_4 - 32h^5a_5 &= -f_n + f_{n-2} + 2hf'_n \\
-4ha_2 + 12h^2a_3 - 32h^3a_4 + 80h^4a_5 &= -f'_n + f'_{n-2}. \tag{2.20}
\end{aligned}$$

Using Gaussian elimination on the last four equations in (2.20) gives

$$\begin{aligned}
h^2a_2 - h^3a_3 + h^4a_4 - h^5a_5 &= -f_n + f_{n-1} + hf'_n \\
h^2a_3 - 2h^3a_4 + 3h^4a_5 &= -\frac{2}{h}(f_n - f_{n-1}) + f'_n + f'_{n-1} \\
4h^4a_4 - 16h^5a_5 &= -5f_n + 4f_{n-1} + f_{n-2} + 2hf'_n + 4hf'_{n-1} \\
4h^4a_5 &= \frac{3}{h}(-f_n + f_{n-2}) + f'_n + 4f'_{n-1} + f'_{n-2}. \tag{2.21}
\end{aligned}$$

From backward substitution, the coefficients  $a_2, a_3, a_4$ , and  $a_5$  are

$$\begin{aligned}
a_2 &= \frac{1}{4h^2} \{-23f_n + 16f_{n-1} + 7f_{n-2} + h[12f'_n + 16f'_{n-1} + 2f'_{n-2}]\} \\
a_3 &= \frac{1}{4h^3} \{-33f_n + 16f_{n-1} + 17f_{n-2} + h[13f'_n + 32f'_{n-1} + 5f'_{n-2}]\} \\
a_4 &= \frac{1}{4h^4} \{-17f_n + 4f_{n-1} + 13f_{n-2} + h[6f'_n + 20f'_{n-1} + 4f'_{n-2}]\} \\
a_5 &= \frac{1}{4h^5} \{-3f_n + 3f_{n-2} + h[f'_n + 4f'_{n-1} + f'_{n-2}]\}. \tag{2.22}
\end{aligned}$$

Inserting the coefficients from (2.22) into the generic explicit compact method (2.16) generates the three step explicit compact scheme

$$y_{n+1} = y_n + \frac{h}{240} \{-949f_n + 608f_{n-1} + 581f_{n-2} + h[637f'_n + 1080f'_{n-1} + 173f'_{n-2}]\}. \tag{2.23}$$

### 2.3 Implicit Compact Method

The formula for the two step implicit compact scheme is derived in this section. Coefficient equation (2.12) must be used to obtain these expressions, and the shift in the time coordinate used for explicit methods is applied here as well.

The time coordinate shift changes (2.12) to

$$\begin{aligned}
 a_0 &= f_n \\
 a_1 &= f'_n \\
 h^2 a_2 - h^3 a_3 + h^4 a_4 - h^5 a_5 &= -f_n + f_{n-1} + hf'_n \\
 -2ha_2 + 3h^2 a_3 - 4h^3 a_4 + 5h^4 a_5 &= -f'_n + f'_{n-1} \\
 h^2 a_2 + h^3 a_3 + h^4 a_4 + h^5 a_5 &= f_{n+1} - f_n - hf'_n \\
 2ha_2 + 3h^2 a_3 + 4h^3 a_4 + 5h^4 a_5 &= f'_{n+1} - f'_n.
 \end{aligned} \tag{2.24}$$

Gaussian elimination reduces the last four equations in (2.24) to

$$\begin{aligned}
 h^2 a_2 - h^3 a_3 + h^4 a_4 - h^5 a_5 &= -f_n + f_{n-1} + hf'_n \\
 h^2 a_3 - 2h^3 a_4 + 3h^4 a_5 &= \frac{2}{h}(-f_n + f_{n-1}) + f'_n + f'_{n-1} \\
 4h^4 a_4 - 4h^5 a_5 &= f_{n+1} + 4f_n - 5f_{n-1} - h(4f'_n + 2f'_{n-1}) \\
 4h^4 a_5 &= \frac{3}{h}(-f_{n+1} + f_{n-1}) + f'_{n+1} + 4f'_n + f'_{n-1}.
 \end{aligned} \tag{2.25}$$

Backward substitution gives the coefficients  $a_2, a_3, a_4,$  and  $a_5$  for the implicit two step formula

$$\begin{aligned}
 a_2 &= \frac{1}{4h^2} \{4f_{n+1} - 8f_n + 4f_{n-1} + h[-f'_{n+1} + f'_{n-1}]\} \\
 a_3 &= \frac{1}{4h^3} \{5f_{n+1} - 5f_{n-1} - h[f'_{n+1} + 8f'_n + f'_{n-1}]\}
 \end{aligned}$$

$$a_4 = \frac{1}{4h^4} \{-2f_{n+1} + 4f_n - 2f_{n-1} + h[f'_{n+1} - f'_{n-1}]\}$$

$$a_5 = \frac{1}{4h^5} \{-3f_{n+1} + 3f_{n-1} + h[f'_{n+1} + 4f'_n + f'_{n-1}]\}. \quad (2.26)$$

The approximate solution to the first order differential equation (2.1) can now be calculated from (2.14)

$$y(t_{i+1}) = y(t_i) + \int_0^h (a_0 + a_1t + a_2t^2 + \Lambda + a_{2m+1}t^{2m+1})dt ,$$

$$y_{n+1} = y_n + ha_0 + \frac{h^2}{2}a_1 + \Lambda + \frac{h^{2m+2}}{2m+2}a_{2m+1}. \quad (2.27)$$

Using the coefficients  $a_0, a_1, a_2, a_3, a_4,$  and  $a_5$  in (2.27) gives the two step implicit compact equation

$$y_{n+1} = y_n + \frac{h}{240} \{101f_{n+1} + 128f_n + 11f_{n-1} + h[-13f'_{n+1} + 40f'_n + 3f'_{n-1}]\}. \quad (2.28)$$

Since equation (2.28) generally has the unknown  $y_{n+1}$  on the right hand side in the  $f_{n+1}$  and  $f'_{n+1}$  terms, it is combined with some explicit method in a predictor-corrector scheme or it is used with a Newton's method solution. These complications are justified by the larger stability region of the implicit method, as discussed in the next chapter.

#### 2.4 Truncation Error

The local truncation error for a numerical method is defined as the difference between the exact solution at  $t_{n+1}$  and the numerical method evaluated with the exact solution  $y(t_{n+1})$ , all divided by the grid spacing  $h$  (Gear, 1971). The order of a method is defined as the maximum local truncation error over the interval  $[a, b]$  on which the problem is defined for a given  $h$ . Hence the integral of the last term in (2.6) divided by  $h$  will give the order of the method.

For explicit compact methods the equation (2.6) is

$$f(t, y(t)) = H_{2m-1}(t) + \frac{(t-t_n)^2(t-t_{n-1})^2 \Lambda(t-t_{n+1-m})^2}{(2m)!} f^{(2m)}(\xi(t)). \quad (2.29)$$

The local truncation error is therefore

$$\tau_{n+1} = \frac{1}{h(2m)!} \int_{t_n}^{t_{n+1}} (t-t_n)^2(t-t_{n-1})^2 \Lambda(t-t_{n+1-m})^2 f^{(2m)}(\xi(t)) dt. \quad (2.30)$$

Changing the variable by substituting  $t = t_n + sh$ , with  $s \in (0,1)$  (Gear, 1971)

$$\tau_{n+1} = \frac{h^{2m}}{(2m)!} \int_0^1 s^2(s+1)^2 \Lambda(s+m-1)^2 f^{(2m)}(\xi_n) ds. \quad (2.31)$$

Since the function  $s^2(s+1)^2 \Lambda(s+m-1)^2$  does not change sign on  $[0,1]$ , the Weighted Mean Value Theorem for integrals states that for some  $\mu_n$  such that  $t_{n+1-m} < \mu_n < t_{n+1}$

$$\tau_{n+1} = \frac{h^{2m} f^{(2m)}(\mu_n)}{(2m)!} \int_0^1 s^2(s+1)^2 \Lambda(s+m-1)^2 ds. \quad (2.32)$$

For the two step explicit scheme (2.19), the local truncation error is

$$\tau_{n+1} = \frac{31}{720} h^4 y^{(5)}(\mu_n), \quad (2.33)$$

and for the three step explicit scheme the local truncation error is (Thompson and Liu, 2004)

$$\tau_{n+1} = \frac{53}{4725} h^6 y^{(7)}(\mu_n). \quad (2.34)$$

For the implicit case, the Hermite interpolation formula (2.6) gives

$$f(t, y(t)) = H_{2m+1}(t) + \frac{(t-t_{n+1})^2(t-t_n)^2 \Lambda(t-t_{n+1-m})^2}{(2m+2)!} f^{(2m+2)}(\xi(t)). \quad (2.35)$$

Using the same approach as in the explicit case, the expression for the implicit local truncation error is

$$\tau_{n+1} = \frac{h^{2m+2} f^{(2m+2)}(\mu_n)}{(2m+2)!} \int_0^1 (s-1)^2(s)^2(s+1)^2 \Lambda(s+m-1)^2 ds, \quad (2.36)$$

for  $\mu_n$  such that  $t_{n+1-m} < \mu_n < t_{n+1}$ . The truncation error for the two step implicit compact method equation (2.28) is (Thompson and Liu, 2004)

$$\tau_{n+1} = \frac{1}{9450} h^6 y^{(7)}(\mu_n). \quad (2.37)$$

One advantage common to all second derivative methods is smaller error coefficients (Lambert, 1991). The fourth order explicit Adams-Bashforth coefficient is over eight times larger than the fourth order explicit compact coefficient, the five step sixth order Adams-Bashforth coefficient is over sixty eight times greater than the sixth order explicit compact coefficient, and the sixth order Adams-Moulton coefficient is more than two orders of magnitude greater than that of the two step implicit sixth order compact method.

## CHAPTER 3

### CONSISTENCY, CONVERGENCE, AND STABILITY

This chapter discusses consistency, convergence, and stability for the compact methods of the previous chapter. Necessary and sufficient conditions for both consistency and speed of convergence for general second derivative multistep methods of the form

$$y_{n+1} = \sum_{j=0}^{m-1} a_j y_{n-j} + h \sum_{j=-1}^{m-1} b_j y'_{n-j} + h^2 \sum_{j=-1}^{m-1} c_j y''_{n-j} \quad (3.1)$$

are given in section 3.1. A theorem on convergence of numerical methods of the form of (3.1) with  $a_0 = 1$ , and the remaining  $a_j = 0$  for all  $j \in \{1, 2, \dots, m-1\}$  is proved in section 3.2, and section 3.3 discusses stability for the compact methods.

#### 3.1 Consistency and Speed of Convergence

The general concept of consistency is that the difference equation approaches the differential equation as the grid spacing  $h$  goes to zero, which says that the truncation error goes to zero as  $h$  goes to zero. The following definitions will make this general concept more precise.

**Definition 3.1.1** The truncation error  $T_n(y)$  is the difference between  $y(t_{n+1})$  and the method (3.1) evaluated at the exact solution (Atkinson, 1989)

$$T_n(y) = y(t_{n+1}) - \left[ \sum_{j=0}^{m-1} a_j y(t_{n-j}) + h \sum_{j=-1}^{m-1} b_j y'(t_{n-j}) + h^2 \sum_{j=-1}^{m-1} c_j y''(t_{n-j}) \right], \text{ for } n \geq m-1. \quad (3.2)$$

**Definition 3.1.2** The local truncation error for the method (3.1) is defined as

$$\tau_n(y) = \frac{T_n(y)}{h}. \quad (3.3)$$

**Definition 3.1.3** The consistency condition for the numerical method (3.1) is defined by (Atkinson, 1989)



$$\tau(h) \equiv \text{Max}_{m-1 \leq n \leq N} |\tau_n(y)| \rightarrow 0 \text{ as } h \rightarrow 0. \quad (3.4)$$

The rate at which the numerical solution converges to the exact answer is determined by the conditions such that

$$\tau(h) = O(h^k). \quad (3.5)$$

Definition 3.1.4 A multistep method is consistent if (Burden and Faires, 2005)

$$\lim_{h \rightarrow 0} |\tau_n(h)| = 0, \text{ for all } n \in \{m, m+1, K, N\}$$

and

$$\lim_{h \rightarrow 0} |\tilde{y}_n - y(t_n)| = 0, \text{ for all } n \in \{1, 2, K, m-1\}. \quad (3.6)$$

Here the initial condition sets the error at  $n = 0$  to zero, the second of equations (3.6) states that the numerical method used to calculate each  $\tilde{y}_n$  of the remaining initial  $m$  steps must also be consistent for the multistep method to be consistent. Atkinson proves the following theorem for methods where  $c_j = 0$  for all  $j \in \{-1, 0, 1, K, m-1\}$ . The proof for second derivative methods here is similar to that of Atkinson.

Theorem 3.1.1 For a given integer  $k$  the method (3.1) is consistent if and only if

$$\sum_{j=0}^{m-1} a_j = 1 \quad \text{and} \quad -\sum_{j=0}^{m-1} j a_j + \sum_{j=-1}^{m-1} b_j = 1. \quad (3.7)$$

Furthermore, the order condition (3.5) is valid for all functions  $y(t) \in C^{k+1}[a, b]$  if and only if the method (3.1) satisfies (3.7) and

$$\sum_{j=0}^{m-1} (-j)^i a_j + i \sum_{j=-1}^{m-1} (-j)^{i-1} b_j + i(i-1) \sum_{j=-1}^{m-1} (-j)^{i-2} c_j = 1 \quad \text{for all } i \in \{2, 3, K, k\}. \quad (3.8)$$

Proof: Expand  $y(t)$  in a Taylor series about  $t_n$

$$y(t) = \sum_{i=0}^k \frac{1}{i!} (t - t_n)^i y^{(i)}(t_n) + R_{k+1}(t). \quad (3.9)$$

Since  $T_n$  is linear

$$T_n(y(t)) = \sum_{i=0}^k \frac{1}{i!} y^{(i)}(t_n) T_n((t-t)^i) + T_n(R_{k+1}(t)). \quad (3.10)$$

From equation (3.2), evaluating  $T_n((t-t_n)^i)$  for  $i = 0$  gives

$$T_n(1) \equiv \alpha_0 = 1 - \sum_{j=0}^{m-1} a_j. \quad (3.11)$$

For  $i \geq 1$ ,

$$T_n((t-t_n)^i) = (t_{n+1} - t_n)^i - \sum_{j=0}^{m-1} a_j (t_{n-j} - t_n)^i - \sum_{j=-1}^{m-1} \{hb_j i (t_{n-j} - t_n)^{i-1} - h^2 c_j i(i-1) (t_{n-j} - t_n)^{i-2}\}$$

or

$$T_n((t-t_n)^i) = \alpha_i h^i, \quad (3.12)$$

where

$$\alpha_i = 1 - \sum_{j=0}^{m-1} (-j)^i a_j - i \sum_{j=-1}^{m-1} (-j)^{i-1} b_j - i(i-1) \sum_{j=-1}^{m-1} (-j)^{i-2} c_j. \quad (3.13)$$

Therefore

$$T_n(y(t)) = \sum_{i=0}^k \frac{\alpha_i h^i}{i!} y^{(i)}(t_n) T_n((t-t)^i) + T_n(R_{k+1}(t)). \quad (3.14)$$

The reminder term can be rewritten as

$$R_{k+1}(t) = \frac{1}{(k+1)!} (t-t_n)^{k+1} y^{(k+1)}(t_n) + \Lambda, \quad (3.15)$$

hence

$$T_n(R_{k+1}(t)) = \frac{\alpha_{k+1} h^{k+1}}{(k+1)!} y^{(k+1)}(t_n) + O(h^{k+2}), \quad (3.16)$$

as long as  $y(t) \in C^{k+2}[a, b]$ . For the numerical method (3.1) to be consistent, it is necessary for

$\tau(h) = O(h)$ , which requires  $T_n(y) = O(h^2)$ . This means that  $k = 1$  in (3.16), and therefore  $\alpha_0$  and  $\alpha_1$  in equation (3.14) must be zero. Equation (3.11) gives

$$\sum_{j=0}^{m-1} a_j = 1, \quad (3.17)$$

and (3.13) with  $i = 1$  gives

$$-\sum_{j=0}^{m-1} j a_j + i \sum_{j=-1}^{m-1} b_j = 1. \quad (3.18)$$

The first part of Theorem 3.1.1 is therefore proved. If  $\tau(h) = O(h^k)$ , then  $T_n(y) = O(h^{k+1})$ , and equations (3.16) and (3.17) show that this is true if and only if  $\alpha_i \equiv 0$  for all  $i \in \{2, 3, \dots, k\}$ . Therefore

$$\sum_{j=0}^{m-1} (-j)^i a_j + i \sum_{j=-1}^{m-1} (-j)^{i-1} b_j + i(i-1) \sum_{j=-1}^{m-1} (-j)^{i-2} c_j = 1, \quad (3.19)$$

for all  $i \in \{2, 3, \dots, k\}$ , and the theorem is proved.

Since  $a_0 = 1$ , and  $a_j = 0$  if  $j \in \{1, 2, \dots, m-1\}$  for the compact methods, the first equation in the consistency condition (3.7) is always satisfied, and the second condition reduces to

$$\sum_{j=-1}^{m-1} b_j = 1. \quad (3.20)$$

The order condition (3.8) for compact methods also simplifies, and is

$$i \sum_{j=-1}^{m-1} (-j)^{i-1} b_j + i(i-1) \sum_{j=-1}^{m-1} (-j)^{i-2} c_j = 1. \quad (3.21)$$

The only non zero  $b_j$  in the two step explicit compact method are  $b_0$  and  $b_1$ ,

$$b_0 = -\frac{1}{2}, \text{ and } b_1 = \frac{3}{2}, \quad (3.22)$$

therefore the consistency condition (3.20) is satisfied for this case. The three step explicit compact method has non zero  $b_j$  components

$$b_0 = -\frac{949}{240}, \quad b_1 = \frac{608}{240}, \quad \text{and} \quad b_2 = \frac{581}{240}, \quad (3.23)$$

and (3.20) is satisfied for the three step case as well. The non zero  $b_j$  for the two step implicit compact method are

$$b_{-1} = \frac{101}{240}, \quad b_0 = \frac{128}{240}, \quad \text{and} \quad b_1 = \frac{11}{240}, \quad (3.24)$$

and this method is also consistent.

The rate of convergence criteria (3.21) is also satisfied for all three compact methods. The truncation errors for the compact methods are defined in Chapter 2, equations (2.33), (2.34), and (2.37) for the explicit two step scheme, the three step explicit scheme, and the implicit two step scheme respectively. The order of each truncation error will be shown to satisfy the rate of convergence criteria (3.21) for all  $i \in \{2, 3, \dots, k\}$ . For the implicit two step method,  $k = 6$ .

$$\begin{aligned} i = 2: & \quad 2(b_{-1} - b_1) + 2(c_{-1} + c_0 + c_1) = 2\left(\frac{101}{240} - \frac{11}{240}\right) + 2\left(-\frac{13}{240} + \frac{40}{240} + \frac{3}{240}\right) = 1 \\ i = 3: & \quad 3(b_{-1} + b_1) + 6(c_{-1} - c_1) = 3\left(\frac{101}{240} + \frac{11}{240}\right) + 6\left(-\frac{13}{240} - \frac{3}{240}\right) = 1 \\ i = 4: & \quad 4(b_{-1} - b_1) + 12(c_{-1} + c_1) = 4\left(\frac{101}{240} - \frac{11}{240}\right) + 12\left(-\frac{13}{240} + \frac{3}{240}\right) = 1 \\ i = 5: & \quad 5(b_{-1} + b_1) + 20(c_{-1} - c_1) = 5\left(\frac{101}{240} + \frac{11}{240}\right) + 20\left(-\frac{13}{240} - \frac{3}{240}\right) = 1 \\ i = 6: & \quad 6(b_{-1} - b_1) + 30(c_{-1} + c_1) = 6\left(\frac{101}{240} - \frac{11}{240}\right) + 30\left(-\frac{13}{240} + \frac{3}{240}\right) = 1 \end{aligned} \quad (3.25)$$

For the explicit three step explicit method,  $k = 6$  as well.

$$\begin{aligned}
i = 2: 2(-b_1 - 2b_2) + 2(c_0 + c_1 + c_2) &= 2\left(-\frac{608}{240} - 2\frac{581}{240}\right) + 2\left(\frac{637}{240} + \frac{1080}{240} + \frac{173}{240}\right) = 1 \\
i = 3: 3(b_1 + 4b_2) + 6(-c_1 - 2c_2) &= 3\left(\frac{608}{240} + 4\frac{581}{240}\right) + 6\left(-\frac{1080}{240} - 2\frac{173}{240}\right) = 1 \\
i = 4: 4(-b_1 - 8b_2) + 12(c_1 + 4c_2) &= 4\left(-\frac{608}{240} - 8\frac{581}{240}\right) + 12\left(\frac{1080}{240} + 4\frac{173}{240}\right) = 1 \\
i = 5: 5(b_1 + 16b_2) + 20(-c_1 - 8c_2) &= 5\left(\frac{608}{240} + 16\frac{581}{240}\right) + 20\left(-\frac{1080}{240} - 8\frac{173}{240}\right) = 1 \\
i = 6: 6(-b_1 - 32b_2) + 30(c_1 + 16c_2) &= 6\left(-\frac{608}{240} - 32\frac{581}{240}\right) + 30\left(\frac{1080}{240} + 16\frac{173}{240}\right) = 1. \quad (3.26)
\end{aligned}$$

Finally, the two step explicit method has  $k = 4$ .

$$\begin{aligned}
i = 2: 2(-b_1) + 2(c_0 + c_1) &= 2\left(-\frac{18}{12}\right) + 2\left(\frac{17}{12} + \frac{7}{12}\right) = 1 \\
i = 3: 3(b_1) + 6(-c_1) &= 3\left(\frac{18}{12}\right) + 6\left(-\frac{7}{12}\right) = 1 \\
i = 4: 4(-b_1) + 12(c_1) &= 4\left(-\frac{18}{12}\right) + 12\left(\frac{7}{12}\right) = 1. \quad (3.27)
\end{aligned}$$

The compact methods are therefore consistent, and both the two step implicit and three step explicit schemes are sixth order, while the two step explicit method is a fourth order.

### 3.2 Convergence

Let  $y(t_n)$  be the exact solution to the first order differential equation (2.1), and define  $y_n$  as the solution to the difference equation associated with some numerical method. If

$$\lim_{h \rightarrow 0} |y(t_n) - y_n| = 0 \text{ for all } n \in \{0, 1, K, N\}, \quad (3.28)$$

then the numerical method is said to be convergent (Gear, 1971). For multistep methods, equation (3.28) must apply to the method used to generate the first  $m - 1$  steps also.

Theorem 3.2.1. Apply the numerical method (3.1) to the first order differential equation

$$y'(t) = f(t, y(t)), \quad y(a) = y_0, \quad \text{and } t \in [a, b]. \quad (3.29)$$

Assume method (3.1) is consistent,  $f(t, y(t))$  and  $f'(t, y(t))$  both satisfy a Lipschitz condition,

$a_j \geq 0$  for all  $j \in \{0, 1, \dots, m-1\}$ , and assume the initial errors satisfy

$$\eta(h) \equiv \text{Max}_{0 \leq j \leq m-1} |y(t_j) - y_j| \rightarrow 0 \quad \text{as } h \rightarrow 0. \quad (3.30)$$

Then the method (3.1) is convergent, and

$$\text{Max}_{a \leq t_j \leq b} |y(t_j) - y_j| \leq \alpha_1 \eta(h) + \alpha_2 \tau(h). \quad (3.31)$$

Proof: This proof follows that given by Atkinson for  $c_k \equiv 0$  for  $k \in \{-1, 0, 1, \dots, m-1\}$ .

Rearranging (3.2) and using  $f(t, y(t))$  for  $y'(t)$ ,

$$y(t_{n+1}) = \sum_{j=0}^{m-1} a_j y(t_{n-j}) + h \sum_{j=1}^{m-1} b_j f(t_{n-j}, y(t_{n-j})) + h^2 \sum_{j=1}^{m-1} c_j f'(t_{n-j}, y(t_{n-j})) + h \tau_n(y). \quad (3.32)$$

Define  $e_k = y(t_k) - y_k$ , and subtract the numerical method (3.1) from (3.32),

$$\begin{aligned} e_{n+1} = & \sum_{j=0}^{m-1} a_j e_{n-j} + h \sum_{j=1}^{m-1} b_j [f(t_{n-j}, y(t_{n-j})) - f(t_{n-j}, y_{n-j})] \\ & + h^2 \sum_{j=1}^{m-1} c_j [f'(t_{n-j}, y(t_{n-j})) - f'(t_{n-j}, y_{n-j})] + h \tau_n(y). \end{aligned} \quad (3.33)$$

Using the Lipschitz conditions for  $f(t, y(t))$  and  $f'(t, y(t))$  gives

$$|e_{n+1}| \leq \sum_{j=0}^{m-1} a_j |e_{n-j}| + h K_1 \sum_{j=1}^{m-1} |b_j| |e_{n-j}| + h^2 K_2 \sum_{j=1}^{m-1} |c_j| |e_{n-j}| + h \tau(h). \quad (3.34)$$

Define the error bound at the node  $n$  by

$$E_n \equiv \text{Max}_{0 \leq i \leq n} |e_i| \quad \text{for each } n \in \{0, 1, \dots, N(h)\}. \quad (3.35)$$

Combining this definition with (3.34),

$$|e_{n+1}| \leq \sum_{j=0}^{m-1} a_j E_n + h K_1 \sum_{j=1}^{m-1} |b_j| E_{n+1} + h^2 K_2 \sum_{j=1}^{m-1} |c_j| E_{n+1} + h \tau(h). \quad (3.36)$$

The consistency condition (3.7) reduces expression (3.36) to

$$|e_{n+1}| \leq E_n + [hK_1 \sum_{j=-1}^{m-1} |b_j| + h^2 K_2 \sum_{j=-1}^{m-1} |c_j|] E_{n+1} + h\tau(h). \quad (3.37)$$

Since the theorem is concerned with the limiting case  $h \rightarrow 0$ , let  $h < 1$ . Therefore

$$h^2 K_2 < hK_2. \quad (3.38)$$

The error equation (3.37) becomes

$$|e_{n+1}| \leq E_n + hcE_{n+1} + h\tau(h), \quad (3.39)$$

where the constant  $c$  is defined as

$$c \equiv K_1 \sum_{j=-1}^{m-1} |b_j| + K_2 \sum_{j=-1}^{m-1} |c_j|. \quad (3.40)$$

The right hand side of (3.39) is also a bound for  $E_n$ , and therefore

$$E_{n+1} \leq E_n + hcE_{n+1} + h\tau(h), \quad (3.41)$$

and hence

$$E_{n+1} \leq \frac{E_n}{1-hc} + \frac{h\tau(h)}{1-hc} \leq (1+2hc)E_n + 2h\tau(h). \quad (3.42)$$

Recursive application of (3.42) to  $E_n$  gives

$$E_n \leq (1+2hc)^n E_0 + [1 + (1+2hc) + \Lambda + (1+2hc)^{n-1}] h\tau(h). \quad (3.43)$$

But

$$(1+2hc)^n \leq e^{2nhc} = e^{2c(t_n-a)} \leq e^{2c(b-a)}, \quad (3.44)$$

and

$$[1 + (1+2hc) + \Lambda + (1+2hc)^{n-1}] = \frac{(1+2hc)^n - 1}{2hc} \leq \frac{e^{2c(b-a)} - 1}{2hc}. \quad (3.45)$$

Using (3.44), (3.45), and the fact that  $E_0 \leq \eta(h)$ ,

$$E_n \leq e^{2c(b-a)}\eta(h) + \left[ \frac{e^{2c(b-a)} - 1}{c} \right] \tau(h), \quad (3.46)$$

and the theorem is proved.

Second derivative methods of the form of (3.1) with  $a_j \geq 0$  for  $j \in \{0, 1, \dots, m-1\}$  are therefore convergent numerical methods; consequently the compact schemes are convergent as well.

### 3.3 Stability

A numerical method is considered to be stable if small perturbations in the initial conditions produce correspondingly small changes in the subsequent approximations (Burden and Faires, 2005). A number of definitions pertinent to multistep methods are also given in this section (Enright, 1974).

Definition 3.3.1. The stability region  $R$  associated with a multistep formula is defined as the set

$$R = \{h\lambda : \text{the formula applied to } y' = \lambda y, y(t_0) = y_0, \text{ with constant step size } h > 0, \\ \text{produces a sequence } \{y_n\} \text{ satisfying } y_n \rightarrow 0 \text{ as } n \rightarrow \infty\}.$$

Definition 3.3.2. A formula is A-stable if the region  $R$  contains the open left half-plane.

Definition 3.3.3. A formula is stiffly-stable if  $R$  contains a region of the form  $R_1 \cup R_2$  where

$$R_1 = \{\operatorname{Re}(z) \leq D < 0\} \\ R_2 = \{D < \operatorname{Re}(z) < 0, |\operatorname{Im}(z)| < \mathcal{G}\}. \quad (3.47)$$

Definition 3.3.4. A formula is stable at infinity if there exist a real number  $\beta < 0$  such that

$$\operatorname{Sup}_{h\lambda < \beta} \left| \frac{y_n}{y_{n-1}} \right| < 1. \quad (3.48)$$

Applying the numerical method (3.1) to the special problem  $y' = \lambda y, y(t_0) = y_0$ , used to investigate stability in multistep methods gives



$$(1 - h\lambda b_{-1} - h^2 \lambda^2 c_{-1})y_{n+1} - \sum_{j=0}^{m-1} (a_j + h\lambda b_j + h^2 \lambda^2 c_j)y_{n-j} = 0. \quad (3.49)$$

This is a homogeneous linear difference equation of order  $m$ . Looking for solutions of the form (Henrici, 1962; Isaacson and Keller, 1966)

$$y_k = r^k, \text{ for } k \geq 0, \quad (3.50)$$

changes (3.49) into

$$(1 - h\lambda b_{-1} - h^2 \lambda^2 c_{-1})r^m - \sum_{j=0}^{m-1} (a_j + h\lambda b_j + h^2 \lambda^2 c_j)r^{m-1-j} = 0, \quad (3.51)$$

after dividing out  $r^{n-m+1}$ . This equation is the characteristic equation, and the left hand side is the characteristic polynomial (Atkinson, 1989). When  $\lambda = 0$ , the characteristic equation reduces to

$$r^m - \sum_{j=0}^{m-1} a_j r^{m-1-j} = 0. \quad (3.52)$$

The root condition associated with (3.52) comes from Burden and Faires.

Definition 3.3.5. Let  $r_1, r_2, \dots, r_m$  be the roots of the characteristic equation (3.52) associated with the difference equation (3.1). These roots are not necessarily distinct. If  $|r^i| \leq 1$  for all  $i \in \{1, 2, \dots, m\}$ , and if all roots with  $|r^i| = 1$  are simple roots, then the difference method satisfies the root condition.

Definition 3.3.6. Numerical methods that satisfy the root condition and have  $|r^i| = 1$  for only one value of  $i$  are called strongly stable.

Definition 3.3.7. Numerical methods that satisfy the root condition and have more than one root of magnitude one are called weakly stable.

Definition 3.3.8. Numerical methods that do not satisfy the root condition are called unstable.

The compact methods and the Adams methods both have  $a_0 = 1$ , and the remaining  $a_j = 0$ . Hence the characteristic equation for compact and Adams methods is

$$r^m - r^{m-1} = r^{m-1}(r - 1) = 0, \quad (3.53)$$

and both methods are therefore strongly stable.

Consider the characteristic equation with non-zero  $\lambda$  for the implicit two step compact method,

$$(1 - h\lambda b_{-1} - h^2\lambda^2 c_{-1})r^2 - (a_0 + h\lambda b_0 + h^2\lambda^2 c_0)r - (h\lambda b_1 + h^2\lambda^2 c_1) = 0. \quad (3.54)$$

The principle root of (3.54) is given by

$$\begin{aligned} r_1 &= \frac{1}{2} \left[ \frac{a_0 + b_0 z + c_0 z^2 + \sqrt{(a_0 + b_0 z + c_0 z^2)^2 + 4(1 - b_{-1} z - c_{-1} z^2)(b_1 z + c_1 z^2)}}{1 - b_{-1} z - c_{-1} z^2} \right], \\ &= \frac{1}{2} \left[ \frac{240 + 128z + 40z^2}{240 - 101z + 13z^2} \right] + \\ &\frac{1}{2} \left[ \frac{\sqrt{(240 + 128z + 40z^2)^2 + 4(240 - 101z + 13z^2)(11z + 3z^2)}}{240 - 101z + 13z^2} \right]. \end{aligned} \quad (3.55)$$

A plot of the stability region for the two step implicit compact method is shown in Figure 3.1. The intersection of the region of stability with the negative real axis is  $-8$  for the two step compact implicit method; whereas the sixth order implicit Adams-Moulton method intersects the negative real axis at approximately  $-1.2$ .

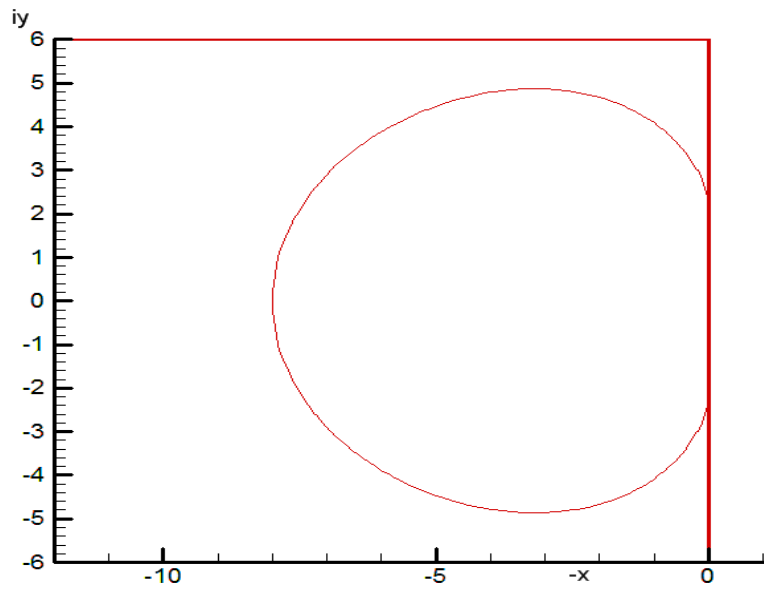


Figure 3.1 Region of stability for the two step implicit compact method. The method is stable for all  $h\lambda$  inside the region.

## CHAPTER 4

### NUMERICAL EXAMPLES

The numerical results of a number of differential equations are presented in this chapter. Section 4.1 compares the sixth order two step compact implicit method with the stiffly-stable sixth order Enright method and the sixth order Adams-Moulton method on three first order differential equations. Section 4.2 applies the compact method to a number of second order differential equations with  $f(t, y(t))$  discontinuous on a set of measure zero. Concluding remarks are in Section 4.3.

#### 4.1 Comparing Compact, Enright, and Adams-Moulton Methods

The two step implicit compact method is compared to two well known methods on three first order test cases, the stiffly-stable sixth order four step implicit Enright method (Enright, 1974)

$$y_{n+1} = y_n + h \left( \frac{3133}{5760} f_{n+1} + \frac{47}{90} f_n - \frac{41}{480} f_{n-1} + \frac{1}{45} f_{n-2} - \frac{17}{5760} f_{n-3} \right) - h^2 \frac{3}{32} f'_{n+1}, \quad (4.1)$$

and the sixth order five step Adams-Moulton scheme with the difference equation

$$y_{n+1} = y_n + \frac{h}{1440} (475 f_{n+1} + 1427 f_n - 798 f_{n-1} + 482 f_{n-2} - 173 f_{n-3} + 27 f_{n-4}). \quad (4.2)$$

Since all three methods are implicit, Newton's method is used to obtain solutions to the difference equations.

The first test differential equation is (Thompson and Liu, 2011)

$$y'(t) = te^{3t} - 2y(t), \quad y(0) = 0, \quad 0 \leq t \leq 1. \quad (4.3)$$

The exact solution for problem (4.3) is

$$y(t) = .2(t - .2)e^{3t} + .04e^{-2t}. \quad (4.4)$$

The problem (4.3) was run in double precision and with grid spacing  $h = 0.1$ . The first five steps of each method are set equal to the exact solution for comparison purposes. The results for the

three methods are shown in Table 4.1. The compact method has approximately one order of magnitude better error than Enright's method, and two orders of magnitude improvement over the Adams-Moulton method.

Table 4.1 Comparison of the absolute value of error for sixth order implicit compact, Enright, and Adams-Moulton methods for problem (4.3).

Time	Exact Solution	Compact Error times E-6	Enright Error times E-6	Adams-Moulton Error times E-6
.5	.28361652	.03693388	.38269513	3.3366422
.6	.49601957	.08243833	.84556844	7.2018678
.7	.82648027	.14063143	1.4402560	12.350526
.8	1.3308570	.21747165	2.2261762	19.060274
.9	2.0897744	.32105130	3.2866842	28.163294
1.0	3.2190993	.46245055	4.7356017	40.598609

The second test problem is a stiff problem (Thompson and Liu, 2011),

$$y'(t) = -20(y - t^2) + 2t, \quad y(0) = \frac{1}{3}, \quad -.4 \leq t \leq 1, \quad (4.5)$$

with exact solution

$$y(t) = t^2 + \frac{1}{3}e^{-20t}. \quad (4.6)$$

The numerical solution for each method was run in double precision with  $h = 0.1$ . The first five steps were set equal to the exact values in order to test each method during the stiff segment of the problem. The results are shown in Table 4.2, and the errors for the compact case are plotted

Table 4.2 Comparison of the absolute value of error for sixth order implicit compact, Enright, and Adams-Moulton methods for problem (4.5).

Time	Exact Solution	Compact Error	Enright Error	Adams-Moulton Error
0.1	.05511176	.0021955527	.113595974	8.9208512
0.2	.04610521	.00093713491	.013323595	4.0705533
0.3	.09082625	.00026894149	.009719635	8.5288475
0.4	.16011182	.64867790 E-4	10.195924 E-4	11.340270
0.5	.25001513	.14201195 E-4	7.6252280 E-4	15.340988
0.6	.36000205	.29261845 E-5	22.279349 E-5	20.879584
0.7	.49000028	.57899321 E-6	99.305093 E-6	28.498281
0.8	.64000004	.11136651 E-6	33.355624 E-6	38.825679
0.9	.81000000	.20989679 E-7	133.51016 E-7	52.900019
1.0	1.0000000	.38975043 E-8	487.89183 E-8	72.088495

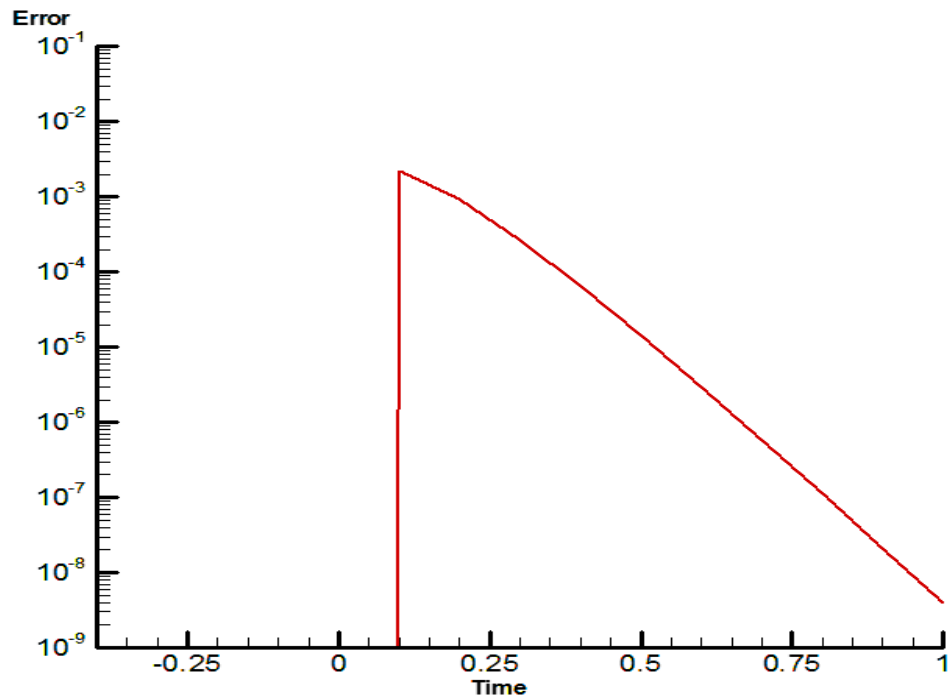


Figure 4.1 Absolute value of the error for the two step compact implicit method applied to problem (4.5)

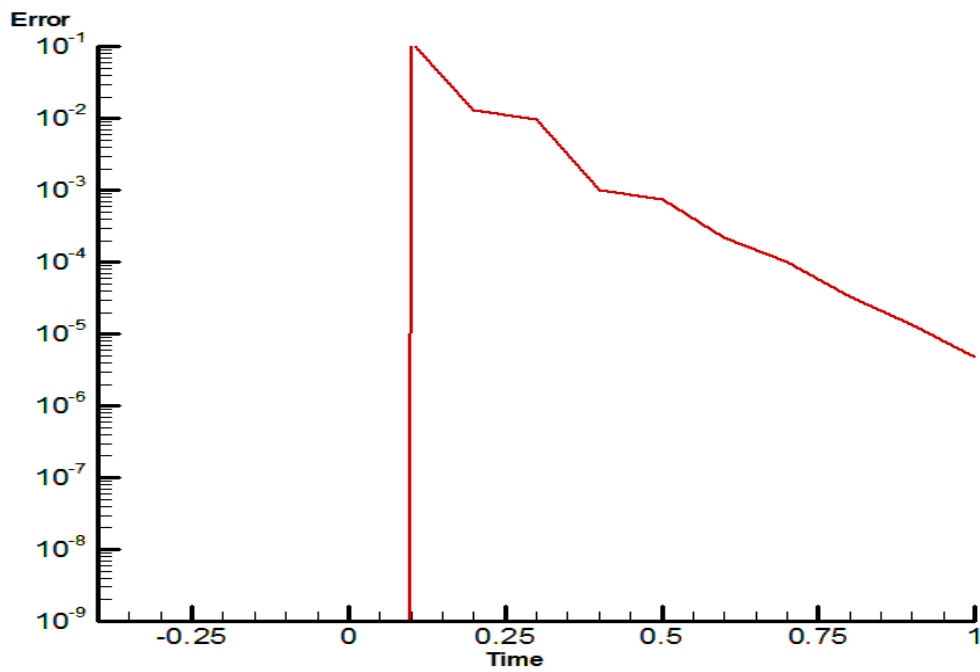


Figure 4.2 Absolute value of the error for the four step Enright method for problem (4.5).

in Figure 4.1, and the error for Enright's method is plotted in Figure 4.2. As Table 4.2 shows, the Adams-Moulton method is unstable for this problem, and hence the error was not plotted for Adams-Moulton. Figures (4.1) and (4.2) both show that the error degrades during the time when the stiff term is predominant, but improves as the  $t^2$  term grows. The two error graphs show that the compact method had smaller error during the stiff portion and also recovered more quickly than Enright's method for problem (4.5).

The last test problem is a simple stiff problem with  $h\lambda = -8$ , which is on the boundary of the stability region for the two step implicit compact method. This is outside the stability region for The Adams-Moulton scheme, and this method was not attempted for problem three. The problem is the differential equation

$$y'(t) = -80y(t), y(0) = \frac{1}{3}, \text{ and } 0 \leq t \leq 1.5. \quad (4.7)$$

The exact solution for (4.7) is

$$y(t) = \frac{1}{3}e^{-80t}. \quad (4.8)$$

The problem was run with grid spacing  $h = 0.1$ . The error for the compact method is shown in Figure 4.3, with the error for Enright shown in Figure 4.4.

The error graphs show two interesting facts. First, both errors are orders of magnitude above the exact value (for  $t = 1$ ,  $y(t) \approx 10^{-36}$ ). The interesting thing is that the compact scheme has smaller error, but is bordering on instability, while the Enright method has more error than the compact method, but as would be expected, the Enright scheme is clearly stable.

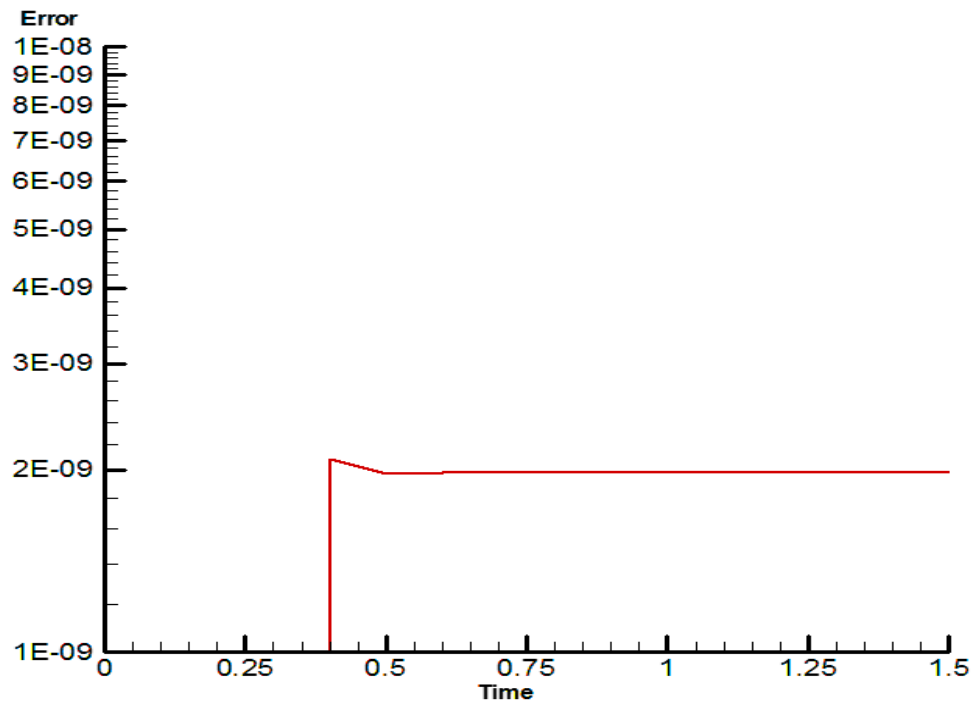


Figure 4.3 Absolute value of the error for the two step implicit compact method for problem (4.7).

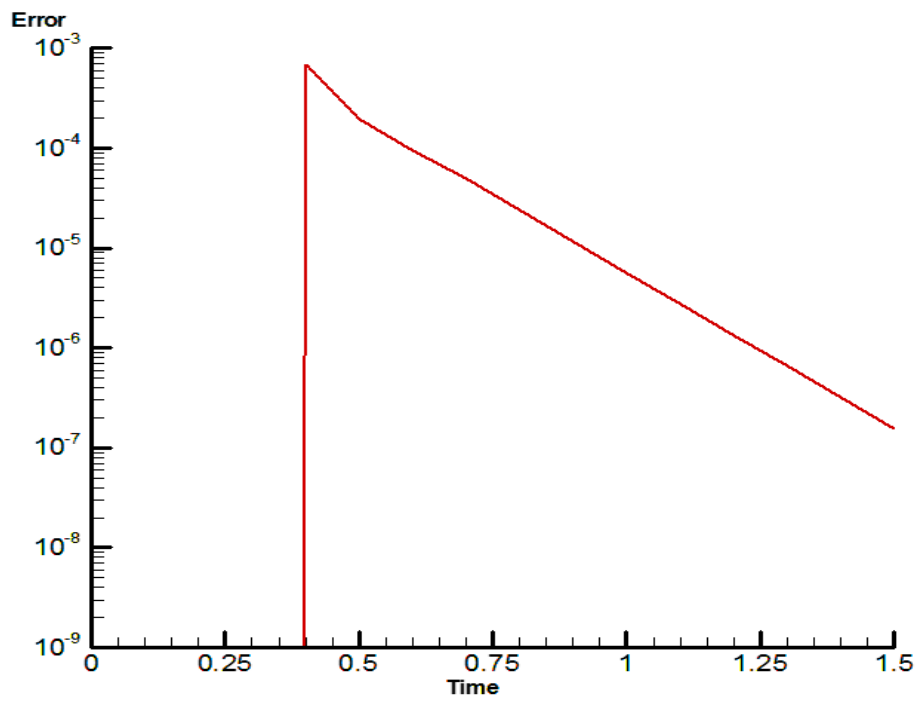


Figure 4.4 Absolute value of the error for Enright's method applied to problem (4.7).



#### 4.2 The Compact Method for Problems with Coulomb Friction

The problems associated with Coulomb friction were discussed in Chapter 1. The approach taken in this work is separation into a sequence of smooth problems, with the end of one segment used as the beginning of a new problem. The initial conditions for the new problem always having a zero first derivative and the position at the time where the velocity goes to zero. This time is generally not a grid point, and thus must be detected by the program. Care must be taken here, since calculating across these discontinuous second derivatives causes big errors, as Figure 1.1 clearly illustrates. This suggest the use of a predictor-corrector scheme since the predictor will be an explicit method, and is thus more suitable for detecting a sign change in the velocity. This sign change is the trigger for the end of segment calculations. This work uses a Hermite polynomial fit to the two grid points preceding the detection of a velocity zero. The Hermite polynomial interpolates the corrector velocity at these two grid points, and Newton's method is used to find the zero of the polynomial. A single step method must be used to find the first  $m - 1$  values of  $y$  and  $y'$ , and a fourth order explicit Runge-Kutta method is used this purpose. The Runge-Kutta method is also used in the crossing point calculations.

The first problem is the unforced spring-block problem described in the first chapter,

$$x'' = -x - .2\text{sgn}(x'), \quad x(0) = 1, \quad x'(0) = 0. \quad (4.9)$$

The solutions are

$$\begin{aligned} x(t) &= .8\cos(t) + .2 & 0 \leq t < \pi \\ x(t) &= .4\cos(t) - .2 & \pi \leq t < 2\pi \\ x(t) &= .2 & t \geq 2\pi, \end{aligned} \quad (4.10)$$

$$\begin{aligned} x'(t) &= -.8\sin(t) & 0 \leq t < \pi \\ x'(t) &= -.4\sin(t) & \pi \leq t < 2\pi \\ x'(t) &= 0 & t \geq 2\pi. \end{aligned} \quad (4.11)$$

The numerical approach used the sixth order explicit three step compact method predictor, with the sixth order implicit two step compact corrector. The numerical results for  $h = 0.1$  and double precision are shown in Figure 4.5. The error plots for (4.9) are presented in Figure 4.6 for  $x(t)$  and Figure 4.7 for  $x'(t)$ .

The last problem has both viscous and Coulomb damping (Taubert, 1976), and is significantly stiff.

$$x'' = -1.6x' - 64x - 2\text{sgn}(x') + 2\cos(\pi t), \quad x(0) = 3.2, \text{ and } x'(0) = 4. \quad (4.12)$$

This calculation used double precision with  $h = 0.005$ . Figure 4.8 is a plot of  $x(t)$ , and Figure 4.9 plots  $x'(t)$ . This problem is dominated by the strong spring force, and the fact that the Coulomb friction and forcing function are about equal. The compact method is in good agreement with Taubert.

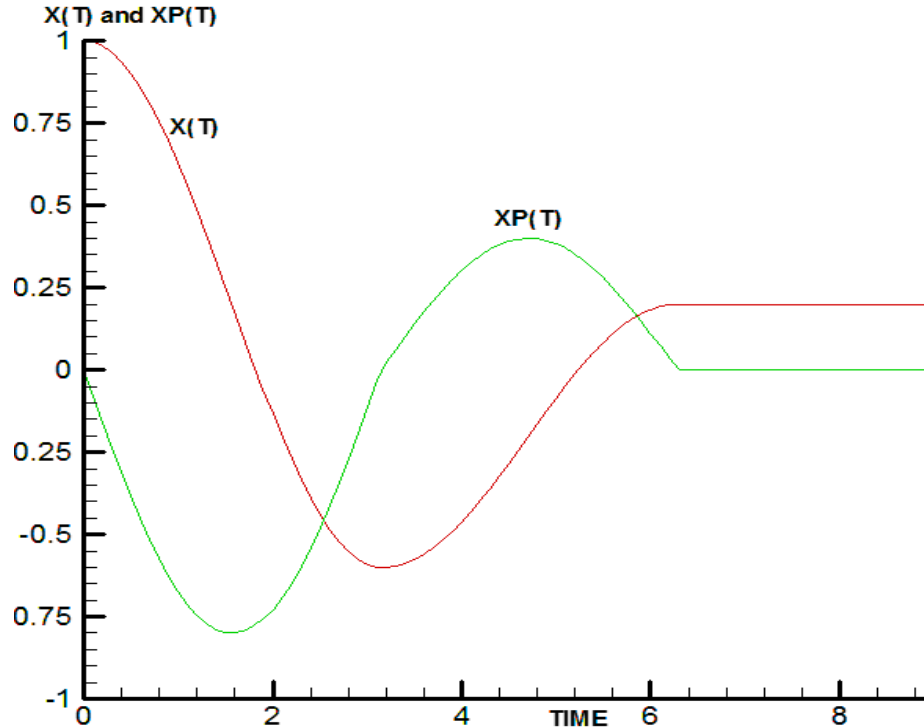


Figure 4.5 Compact method solutions for position and velocity for the unforced Coulomb problem.

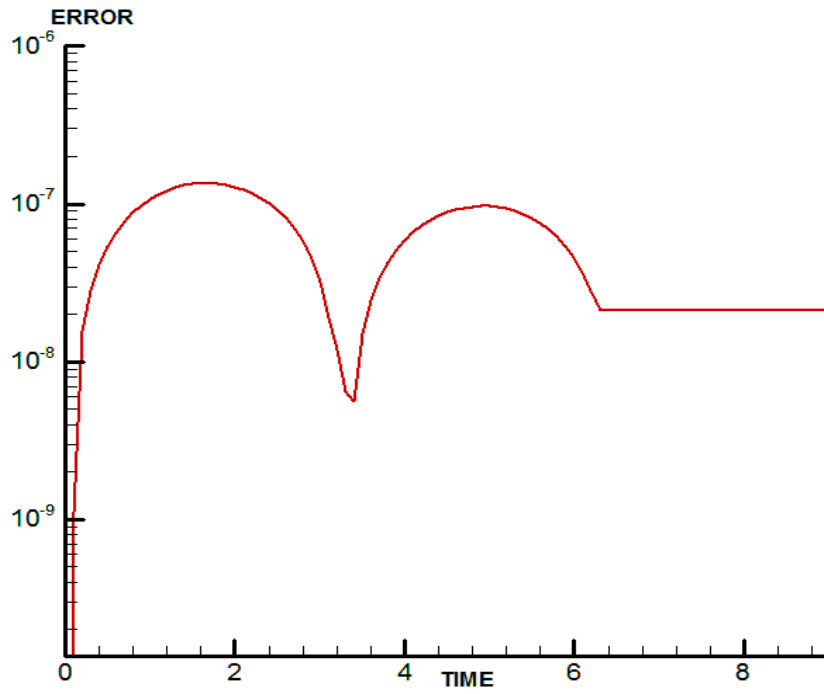


Figure 4.6 Positional error for the unforced Coulomb friction problem.

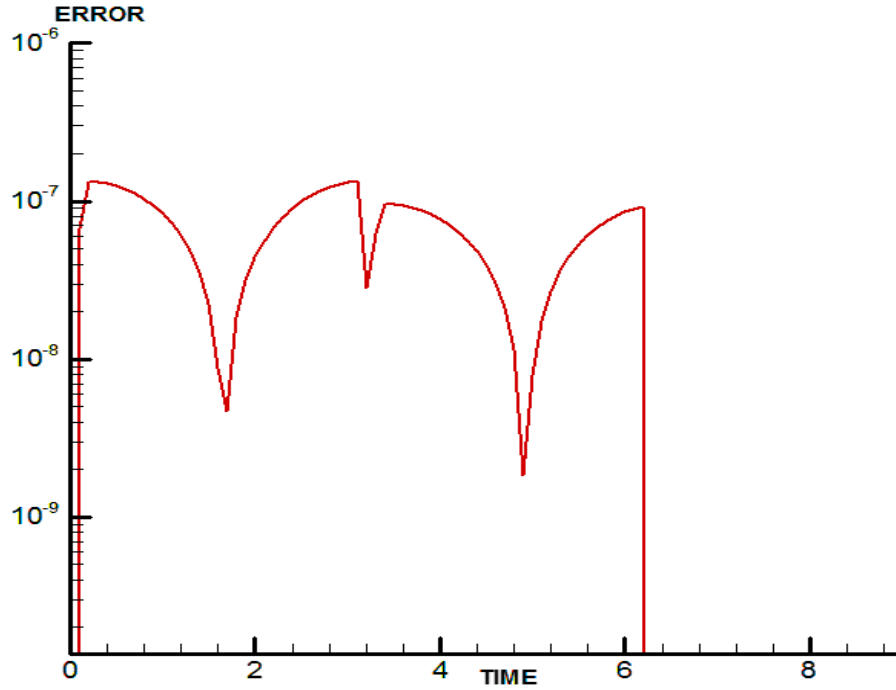


Figure 4.7 Velocity error for the unforced Coulomb friction problem.

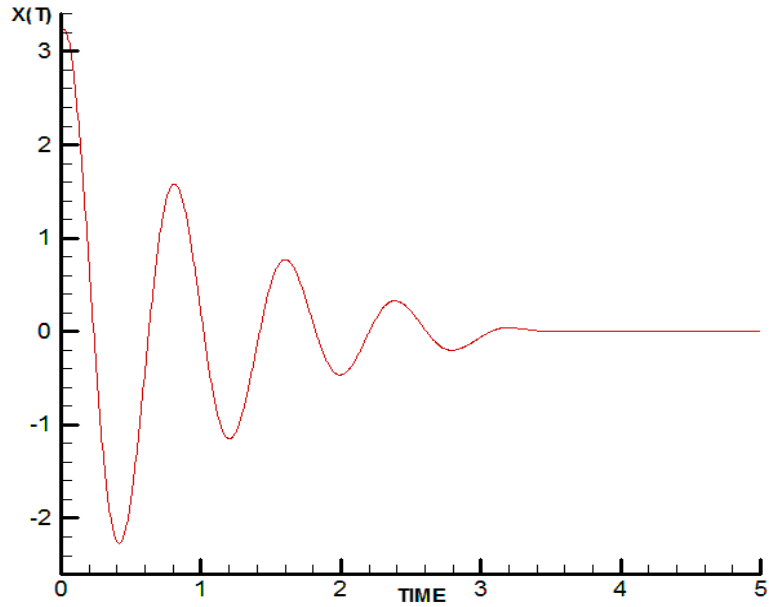


Figure 4.8 Solution for position in problem (4.12) with viscous and Coulomb damping.

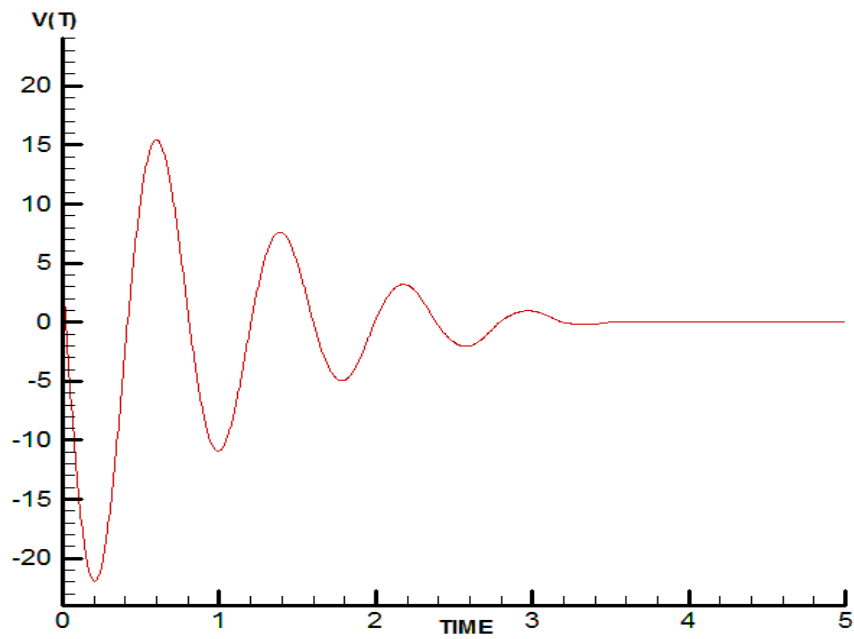


Figure 4.9 Solution for velocity in problem (4.12) with viscous and Coulomb damping.

#### 4.3 Conclusions

The compact method performs well compared to Enright's method and the well known Adams-Moulton method. This new method has order  $2m + 2$ , giving high accuracy for a small

number of steps  $m$ . Although the method is not stiffly-stable, the stability region is larger than that of the Adams methods (even those of smaller order) and explicit Runge-Kutta method. The examples of section 4.1 show that the two step implicit compact method is more accurate than the stiffly-stable four step Enright method within the region of stability of the compact method, and approximately two orders of magnitude more accurate than the five step Adams-Moulton method.

The accuracy of the compact methods combined in a predictor-corrector scheme was used successfully in section 4.2 for both Coloumb and mixed friction problems. The computational difficulty from discontinuity in  $x''(t)$  can be removed by breaking the problem into sections where the acceleration is continuous, and beginning a new initial condition problem at each point of discontinuity. The accuracy of the compact method reduces the error in the calculated values of the crossing points, thereby extending the computational range for these problems.

## REFERENCES

- [1] K. E. Atkinson, *An Introduction to Numerical Analysis*, Wiley, New York, 1989
- [2] R. L. Burden and J. D. Faires, *Numerical Analysis*, Brooks/Cole, Pacific Grove, 2005.
- [3] J. P. Den Hartog, *Mechanical Vibrations*, Dover , New York, 1985.
- [4] W. H. Enright, Derivative Multistep Methods for Stiff Ordinary Differential Equations, SIAM Journal of Numerical Analysis, 11 (1974) 321-331.
- [5] R. P. Feynman, R. B. Leighton, and M. Sands, *Feynman Lectures on Physics*, Addison-Wesley, 1963
- [6] C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, 1971.
- [7] C. W. Gear, Numerical solution of ordinary differential equations: is there anything left to do?, SIAM Review, 23(1981) 10-24.
- [8] P. Henrici, *Elements of Numerical Analysis*, Wiley, New York, 1964.
- [9] D. V. Hutton, *Applied Mechanical Vibrations*, McGraw-Hill, New York, 1981.
- [10] E. Isaacson and H. B. Keller, *Analysis of Numerical Methods*, Wiley, New York, 1966.
- [11] L. S. Jacobsen and R. S. Ayre, *Engineering Vibrations*, McGraw-Hill, New York, 1958.
- [12] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems*, Wiley, Chichester, 1991.
- [13] S. K. Lele, Compact finite difference schemes with spectral-like resolution, Journal of Computational Physics, 103(1992) 16-42.
- [14] N. Obrechhoff, Sur les quadrature mecaniques, Spisanic Bulgar.Akad. Nauk.,65, 191-289.
- [15] L. A. Pipes and L. R. Harvill, *Applied Mathematics for Engineers and Physicist*, McGraw-Hill, New York, 1958.
- [16] L. F. Shampine, Vectorized solution of ODEs in MATLAB with control of residual and error, available at <http://faculty.smu.edu/shampine/current.html>, (2011).

- [17] K. Taubert, Differenzverfahren für schwingungen mit trockener und zäher reibung und für regelungssysteme, Numer. Math. 26 (1976), 379-395.
- [18] J. L. Thompson and C. Liu, Implicit compact scheme for harmonic oscillator equation with Coulomb friction, Proceedings of Dynamic Systems and Applications, 4(2004) 291-296.
- [19] J. L. Thompson and C. Liu, High order implicit compact method for ordinary differential equations, Current Applied Mathematics, Awaiting printing in 2011.
- [20] J. H. Verner, A most efficient Runge-Kutta (13:7,8) pair, available at <http://www.math.sfu.ca/~jverner/>, (2008).

## BIOGRAPHICAL INFORMATION

Mr. Thompson graduated from The University of Texas at Austin in 1968 with a B. S. in Electrical Engineering. He also got a Masters in Physics in 1975, majoring in theoretical physics. Mr Thompson was a staff engineer for the Lockheed-Martin Corporation for the past 29 years, retiring from Lockheed-Martin in January , 2011.