Numerical Methods for the Solution of Fractional Differential Equations

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by Arthur Charles Simpson.

December 2001

Declaration

No part of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other institution of learning, however some parts of the material contained herein have been previously published:

1. Parts of Chapter 12 were presented at the IMACS World Congress on Scientific Computing, Applied Mathematics and Simulation, Lausanne, August 2000.

2. Parts of Chapter 10 have appeared in presentations at the International Conference on Scientific Computing and Mathematical Modelling, Milwaukee, May 2000, at the 19th Biennial Conference on Numerical Analysis, Dundee, June 2001 and in Numerical Algorithms, 26, 333-346, 2001 .

3. Parts of Chapter 10 and Chapter 12 were presented at HERCMA, Athens, September, 2001.

Acknowledgements

I am pleased to acknowledge the help and guidance of my supervisor Professor Neville J. Ford, during the period of work described in this document. I would also like to thank Chester College for its financial support in the form of a Chester College research student bursary.

Abstract

The fractional calculus is a generalisation of the calculus of Newton and Leibniz. The substitution of fractional differential operators in ordinary differential equations substantially increases their modelling power.

Fractional differential operators set exciting new challenges to the computational mathematician because the computational cost of approximating fractional differential operators is of a much higher order than that necessary for approximating the operators of classical calculus.

1. We present a new formulation of the fractional integral.

2. We use this to develop a new method for reducing the computational cost of approximating the solution of a fractional differential equation.

3. This method can be implemented with two levels of sophistication. We compare their rates of convergence, their algorithmic complexity, and their weight set sizes so that an optimal choice, for a particular application, can be made.

4. We show how linear multiterm fractional differential equations can be approximated as systems of fractional differential equations of order at most 1.

Contents

Introduction

0.1 Aim

The aim of this thesis is to demonstrate that numerical approximations to the solutions of fractional differential equations can be obtained with considerably less computational cost than was thought feasible with the algorithms previously available.

0.2 Results

We give a new formulation of the fractional integral as an infinite sum of integrals. From this we develop, as original work, an algorithm, which uses a modification of time valuation, which enables approximate solutions to fractional differential equations to be calculated with a significant saving in computational cost.

We give two versions of our algorithm: firstly one that can be implemented with little programming sophistication but at the sacrifice of some rapidity of convergence, and secondly one that requires more sophisticated programming. These alternatives have different compromises between, accuracy, rates of convergence, and weight set size. The optimal choice therefore depending on the particular application and the end user's objectives.

We show, for the first time, how a multiterm fractional differential equation can be solved as a system of fractional differential equations, of order at most 1, in such a way that the dimensionality of the system is kept to a minimum.

0.3 Content

The theory of calculus presents many problems, particularly in the fields of integral equations and differential equations, whose solution can only be obtain in the form of a numerical approximation. Indeed numerical quadrature, numerical differentiation, and approximation theory have developed conjointly with the abstract theory of the calculus.

The fractional calculus is a generalisation of the calculus of Newton and Leibniz. The substitution of fractional differential operators in ordinary differential equations considerably extends their expressive power and therefore the range of physical, social and economic situations they can or could be used to model.

The increased complexity of fractional differential operators presents exciting new challenges for the computational mathematician to overcome. In particular a considerable penalty, in computational cost, occurs in obtaining an approximate solution when a classical model is generalised to a fractional model since the cost of the problem changes from being linear in time to quadratic in time. This makes it is essential that we develop algorithms that are as efficient as possible.

To keep this thesis self contained we give sufficient abstract theory of the fractional calculus to both enable an understanding of the origin of the

computational cost we need to reduce and to help identify the features of the fractional calculus we can exploit to obtain a reduction in cost.

Traditionally mathematicians have derived their methods for approximating fractional differential operators from consideration of the Riemann-Liouville fractional derivative. More recent papers [12] give their analysis based on the Caputo fractional derivative. The Caputo fractional derivative being the *fractional derivative of choice* amongst mathematical modellers. This is because a fractional differential equation, formulated in terms of the Caputo fractional derivative, utilises the same starting conditions as for an ordinary differential equation of the same bounding integer order.

The methods we will use in numerically approximating the solutions of fractional differential equations are derived from the methods used for solving ordinary differential equations. For this reason we include sufficient theory about the numerics of ordinary differential equations to enable the fractionalised methods to be rigorously derived.

We summarise the standard methods for numerically approximating the solutions of fractional differential equations. We analyse some previously suggested modifications for reducing the cost of implementing these methods and give a new modification, which obtains a substantial reduction in cost without compromising the stability and ultimate convergence of the original method.

We show how the general linear multiterm fractional differential equation can be approximated by using a system of fractional differential equations of order at most 1 and implement this method for several fractional differential equations which have now become standard models.

Chapter 1

We briefly survey the history of the fractional calculus and the mathematical problems it was originally developed to solve.

Chapter 2

We survey some of the modern applications of fractional calculus. This enables us to understand why modelling applications require the use of the fractional derivative as defined by Caputo as opposed to the Riemann-Liouville fractional derivative.

Chapter 3

We give such classical theory of the fractional calculus as we need to understand the cause of the increase of computational effort when calculating numerical approximations. This is the theory most of the existing numerical methods for approximating fractional differential operators are based upon.

Chapter 4

Just as the fractional calculus is a development of the calculus so the numerical methods for fractional differential equations are a development of the numerical methods of ordinary differential equations. We give a selective exposition of numerical ODE theory to establish the concepts we will later use in the numerics of fractional differential equations.

Chapter 5

The truncation of the sum in the Grünwald-Letnikov fractional derivative is the simplest method for approximating a fractional differential operator.

We demonstrate that its convergence properties are insufficient to use it to approximate the solution to the simplest fractional differential equation.

Chapter 6

We give a full exposition of the fractional trapezium rule. We have found this to be the simplest method to implement and its good behaviour has made it suitable to use for prototyping and testing the algorithmic modifications we propose later.

Chapter 7

We give a presentation of the fractional linear multistep method [40, 41, 42, 43] .

Chapter 8

We give our analysis of Podlubny's Finite Memory Principle and an exposition of Sloane and Thomee's Sparse Quadrature Method.

Chapter 9

We give an exposition of the extension of Richardson extrapolation for the trapezium rule to the fractional trapezium rule [17].

Chapter 10

We present a new method of writing the fractional integral as an infinite sum of integrals over the same interval. This enables us to develop a new approach to the distribution of step length in the approximation of the convolution integral in the fractional differential operator.

We do this in two ways which we have chosen to call the *nested mesh* implementation and the *fractal sum* implementation respectively.

In the nested mesh implementation we use a set of nested meshes, similar in concept to the nested subspaces of wavelets, to reduce the computation effort required to approximate the fractional derivative whilst accurately maintaining its history. We show that this reduces the total cost of approximating the fractional derivative to $\mathcal{O}(n \log n)$ with a set of weights of size $\mathcal{O}(\log n)$. This algorithm is simple to implement but this simplicity is achieved at the cost of a slight truncation of the integral and a slight offset in the weights.

In the fractal sum implementation we correct the compromises made in the nested mesh implementation by eliminating the truncation and the offset of the weights. This increases the size of the set of weights to $\mathcal{O}(n)$ but without losing the $\mathcal{O}(n \log n)$ cost of approximating the fractional derivative, however it implementation requires a much more sophisticated computer program.

Chapter11

We give a summary of an extension of fractional linear multistep methods [16] to solving linear multiterm fractional differential equations.

Chapter 12

We show how the standard method used for solving high order ODE's as a system of first order ODE's can by adapted to solve a linear multiterm fractional differential equation and compare this with the approach given in the previous chapter.

Chapter 13

We make some final observations and indicate some future areas of research.

Chapter 1

Fractional Calculus: Its History

1.1 A Brief History of the Calculus

1.1.1 Ancient Greece

The first mention of mathematical ideas that a modern mathematician would identify with the calculus are to be found in the work of Archimedes [9]. In his *method of exhaustion* Archimedes developed a method for estimating areas and volumes of regular geometric objects. In the *method of exhaustion* the area or volume of a regular geometric is shown to be bounded by, what we would regard as, upper and lower Riemann sums which converge to the same limit.

An example of this is that we may estimate the area of a circle by constructing a regular polygon of n sides inside the circle, where the vertices of the polygon lie on the circumference of the circle, and obtain a lower bound for the area of the circle by calculating the area of the polygon by summing the areas of the n isosceles triangles obtained by drawing a line from each vertex of the polygon to the centre of the circle, similarly an upper bound can be obtained by constructing a regular polygon outside the circle whose

sides are tangent to the circumference of the circle.

Therefore we see that Archimedes' *method of exhaustion* is a precursor of definite integration. Due to an inadequate theory of dynamics the Greeks did not have a framework for developing the notion of a derivative.

1.1.2 Newton and Leibniz

With the invention of gunpowder, growing astronomical awareness through the work of Copernicus and Kepler, and the increase of maritime trade it became necessary that mathematicians develop methods for accurately describing and predicting the trajectories of moving objects, to facilitate the accurate targeting of cannons and the prediction of tidal movement.

The differential calculus, as we know it, makes its first appearance in the parallel works of Newton and Leibniz. This calculus is developed through the use of *infinitesimal quantities.* The use of *infinitesimal quantities* continued up to the time of Cauchy before being replaced with the rigorous ϵ-δ method of Weierstrass. Using non-standard models of arithmetic Robinson [52] gave a rigorous development of calculus using *infinitesimals.*

Developed to meet the needs of problems in dynamics, which can be expressed as differential equations, the Newton Leibniz version of calculus places primary significance on the operation of differentiation. Integration theory then playing a secondary role of finding the indefinite integral of a given function. This led to the indefinite integral being interpreted as the anti-derivative, the definite integral taking a subsidiary role, the converse of the early Greek model. Not until the work of Cauchy does the definite integral regain primacy.

1.1.3 Early History - Fractional Calculus

The fractional calculus arises out of a generalisation of the operations of standard calculus. The first suggestion that mathematicians might enquire into the existence of fractional differential operators was by Leibniz who asked what kind of mathematics would statements such as $D^{\frac{1}{2}}$, acting on a function, imply [53].

Euler noted that the formula

$$\frac{d^n x^m}{dx^n} = \frac{m!}{(m-n)!} x^{m-n}, \tag{1.1}$$

could be generalised to

$$\frac{d^q x^p}{dx^q} = \frac{\Gamma(p+1)}{\Gamma(p-q+1)} x^{p-q}. \tag{1.2}$$

This expression interpolates fractional orders of the derivative between integer orders of the derivative but gives no mathematical rationale for so doing. Laplace and Fourier suggested methods whereby their integral transforms could be extended to provide a definition of a fractional derivative.

1.1.4 Abel and the Tautochrone

The solution to the tautochrone problem is reproduced from [45] in Appendix A for ease of reference.

In the problem of the tautochrone we ask what is the shape of a curve C such that the time taken for a frictionless bead to slide down it is independent of the beads starting point, see figure A.1.

In his solution to the problem of the tautochrone Abel [1] was the first person to give a mathematical formulation which would be understood as a

fractional differential operator by later mathematicians. However the fractional differential operator occurs as a mathematical technique not intentionally as part of the mathematical model describing the tautochrone problem.

1.1.5 Abel's Integral Equation and its Solution

When the square root in the denominator of the integral in the tautochrone problem (A.7) is replaced by a root of power μ, where $0 < \mu < 1$, the resulting equation is a known as Abel's integral equation (1.3).

$$\int_0^t \frac{f(s)}{(t-s)^{\mu}} ds = g(t),\ 0 < t \leq,\ 0 < \mu < 1. \tag{1.3}$$

In theorem 1.1 the existence of a unique solution to Abel's integral equation is established. Whilst fractional differential operators are not explicitly mentioned the techniques used in the proof of the theorem , in particular the use of Fubini's theorem [53] to change the order of integration, are fundamental in the development of the fractional calculus.

Theorem 1.1 *([37], 73)*

If $g(t)$ is continuous in $0 < t \leq T$ and

$$\lim_{t \to 0} t^{\alpha} g(t) = C,$$

where $C \neq 0$ and $0 < \alpha < \mu$, then the simple Abel equation has the solution

$$f(t) = \frac{\sin \mu\pi}{\pi} \frac{d}{dt} \int_0^t \frac{g(s)}{(t-s)^{1-\mu}} ds,\ 0 < t \leq T. \tag{1.4}$$

This solution is continuous in $0 < t \leq T$, and satisfies

$$f(t) = (C + o(1)) \frac{\Gamma(1-\alpha)}{\Gamma(1-\mu)\Gamma(\mu-\alpha)} t^{\mu-\alpha-1}, \tag{1.5}$$

as $t \to 0$. Furthermore, this solution is unique in the class of functions of the form $f(t) = t^{\beta} F(t)$, where $\beta > -1$ and $F(t)$ is continuous.

Proof Suppose Abel's integral equation (1.3) has a solution in the class of functions stated then multiply (1.3) by $(x-t)^{\mu-1}$ and integrate from 0 to x. Then

$$\int_0^x \int_0^t \frac{f(s)}{(x-t)^{1-\mu}(t-s)^{\mu}} ds dt = \int_0^x \frac{g(t)}{(x-t)^{1-\mu}} dt. \tag{1.6}$$

Since $f(s)$ is assumed to be of the form $s^{\beta}F(s)$, with $\beta > -1$ and $F(s)$ continuous, we can apply Fubini's theorem, Theorem 3.1, to (1.6) to obtain

$$\int_0^x \int_s^x \frac{dt}{(x-t)^{1-\mu}(t-s)^{\mu}} f(s) ds = \int_0^x \frac{g(t)}{(x-t)^{1-\mu}} dt. \tag{1.7}$$

But

$$\begin{aligned} \int_s^x \frac{dt}{(x-t)^{1-\mu}(t-s)^{\mu}} &= \int_0^{t-s} \frac{du}{(x-s-\mu)^{1-\mu}u^{\mu}} \\ &= \int_0^1 (1-w)^{\mu-1} w^{-\mu} dw \\ &= \Gamma(\mu)\Gamma(1-\mu) \\ &= \frac{\pi}{\sin \mu\pi}. \end{aligned} \tag{1.8}$$

Thus we have

$$\frac{\pi}{\sin \mu\pi} \int_0^x f(s) ds = \int_0^x \frac{g(t)}{(x-t)^{1-\mu}} dt. \tag{1.9}$$

Now when $x \to 0$ so also must $t \to 0$, therefore as $x \to 0$ we have

$$\begin{aligned} \int_0^x \frac{g(t)}{(x-t)^{1-\mu}} dt &\to C \int_0^x \frac{dt}{(x-t)^{1-\mu} t^{\alpha}} \\ &= C \frac{\Gamma(\mu)\Gamma(1-\alpha)}{\Gamma(1+\mu-\alpha)} x^{\mu-\alpha}. \end{aligned} \tag{1.10}$$

Hence the integral can be differentiated everywhere except at $x = 0$. The results (1.4) and (1.5) then follow immediately. Existence is proved by substituting the expression for $f(t)$ into Abel's equation and verifying that it solves (1.3).

1.1.6 The Fractional Integral

Credit for the first mathematically satisfactory formulation of a fractional differential operator goes to Liouville [53]. Liouville considered defining fractional differential operators in several different ways. These were integral formulations, difference equations, and expansions in exponential functions. We will rely primarily on integral formulation in our exposition.

The first person to use the Abel integral to define the fractional integral was Riemann [51], who gives the same definition as we will use, namely:

$$I^{\alpha}\phi = \frac{1}{\Gamma(\alpha)} \int_{o}^{t} \frac{\phi(s)}{(t-s)^{1-\alpha}} ds. \tag{1.11}$$

Grünwald and Letnikov [53] defined a fractional derivative by means of a fractional difference quotient

$$D^{\alpha} f(x) = \lim_{h \to 0} \frac{(\Delta_h^{\alpha} f)(x)}{h^{\alpha}}, \tag{1.12}$$

and Letnikov was able to show, with an appropriate interpretation of the fractional difference, that this definition was equivalent to Riemann's definition of the fractional integral.

Various alternative forms of fractional integrals and derivatives have been proposed. In this thesis we are interested in calculating numerical approximations to the solutions of those fractional differential equations which occur in mathematical models arising out of material science. For this reason we limit our survey of fractional differential operators to the Riemann-Liouville fractional integral and its extension to a Riemann-Liouville fractional derivative and the Caputo fractional derivative by combination with the operation of integer order differentiation.

Chapter 2

Fractional Calculus: Some Applications

2.1 Introduction

The application of the fractional calculus has two main branches mathematical and modelling.

In mathematical applications the fractional calculus is used to obtain the solution to various kinds of integral equations. We give a very brief survey of such applications since problems of this kind have been known for a long time.

In modelling applications the fractional derivative of a function is used to specify the rate of propagation of some aspect of an evolving system. Usually how some parameter of the system varies in time. This is largely a modern development and meeting the computational challenges that this mathematical activity produces is the motivation behind the new results presented in this thesis.

2.2 Mathematical Applications

The fractional calculus can be used for the solution of dual integral equations of the form

$$\int_0^\infty t^{-2\alpha}[1+R(t)]\Psi(t)J_\mu(xt)dt = F(x),\ 0<c<1$$

$$\int_0^\infty t^{-2\beta}\Psi(t)J_\nu(xt)dt = G(x),\ 1<x<\infty,$$

where $R(x)$, $F(x)$ and $G(x)$ are given functions and $\Psi(t)$ is an unknown one.

Equations of this form may occur in mixed boundary value problems in mathematical physics when using the Hankel transform. Extensive examples of the application of fractional calculus to the solution of integral equations are given in [53].

Many special functions of mathematical analysis can by expressed in terms of fractional integrals or derivatives of a smaller set of special functions [46].

2.3 Modelling Applications

The fractional integral arises in a number of nineteenth century mathematical models. In these models a square root of the form $(L-r)^{1/2}$ occurs somewhere in the formulation, or reformulation, of the problem. Here L is a fixed distance and r varies between 0 and L. When this kind of formulation can be expressed in terms of a fractional integral, its solution can then be found by fractionally differentiating both sides of a relationship. An example of this technique was given in the solution to the tautochrone. This technique also occurs in such problems as Liouville's potential problem (on

the relative attraction of two wires to a charged mass), and the weir notch problem (where we require the shape of the notch in a dam such that the rate of flow of water through the notch is a specified function of height) [45].

More recently fractional derivatives have been used to generalise standard differential equations and to produce new equations to model physical processes not adequately described by models using integer order derivatives. In the next three subsections we give examples of situations where fractional differential equations, whose solutions we may wish to approximate, occur.

2.3.1 Linear Viscoelasticity

A material which exhibits elastic and viscous properties is called viscoelastic. There are various standard models of viscoelastic materials. A brief review of the basic elements of classical viscoelastic theory is given in Appendix B. To construct a model of viscoelasticity springs (ideal elastic elements) and dashpots (ideal viscous elements) are combined in series or parallel. Several such elements may be further combined together, in a network, to produce more sophisticated models.

In an ideal elastic element, or spring, the stress $\sigma(t)$ is proportional to the strain $\kappa(t)$ thus

$$\sigma(t) = m\kappa(t). \tag{2.1}$$

In an ideal viscous element, or dashpot, the stress is proportional to the rate of extension thus

$$\sigma(t) = b\frac{d\kappa}{dt}. \tag{2.2}$$

The two simplest models of viscoelasticity are firstly the Voigt model, which comprises a spring and dashpot in parallel, its stress strain relation is given

by

$$\sigma(t) = m\kappa(t) + b\frac{d\kappa}{dt}, \tag{2.3}$$

where $\tau_\kappa = b/m$ is called the retardation time, and secondly the Maxwell model, which comprises a spring and dashpot in series, its stress strain relation is given by

$$\sigma(t) + a\frac{d\sigma}{dt} = m\kappa(t), \tag{2.4}$$

where $\tau_\sigma = a$ is called the relaxation time.

More complicated models, which either add a spring in series to the Voigt model or in parallel to the Maxwell model were introduced by Zener. Such a model is referred to as a standard linear solid or S.L.S.. The stress strain relationship for more complicated models of viscoelasticity take the form of

$$\left[1 + \sum_{k=1}^{p} a_k \frac{d^k}{dt^k}\right] \sigma(t) = \left[m + \sum_{k=1}^{q} b_k \frac{d^k}{dt^k}\right] \kappa(t), \tag{2.5}$$

where $p = q$ or $p = q + 1$.

Following experimental work, which exhibited power-law like creep as opposed to exponential-law creep of the standard models of viscoelasticity, Scott-Blair [54] proposed that the Newtonian dashpot be replaced by an element whose stress stain relationship is modelled by the fractional differential equation

$$\sigma(t) = b\frac{d^\alpha \kappa(t)}{dt^\alpha}. \tag{2.6}$$

Caputo and Mainardi [44] used this suggestion to generalise the S.L.S. model to produce the fractional S.L.S. with stress-strain relationship

$$\left[1 + a\frac{d^\alpha}{dt^\alpha}\right] \sigma(t) = \left[m + b\frac{d^\alpha}{dt^\alpha}\right] \kappa(t), \ 0 < \alpha \leq 1. \tag{2.7}$$

The long term behaviour of solutions to equations of the form (2.7) exhibit the same power-law characteristics as observed by Scott-Blair. This power-law decay is slower than the exponential decay seen in such phenomena as the half-life of radioactive isotopes. Processes exhibiting this power-law behaviour are sometimes referred to as ultraslow processes.

2.3.2 The Basset Force

Here we are concerned with the dynamics of a sphere immersed in an incompressible viscous fluid where the low Reynolds number limit is assumed so that the Navier-Stokes equations describing the fluid motion may be linearised. In this case we assume that the fluid is quiescent and that the sphere begins to move under the action of gravity as may occur when a pea sinks in a glass of crème de menthe. The equation of motion of the sphere can then be expressed in the form

$$\frac{dV}{dt} = -\frac{1}{\sigma_\epsilon}\left[1 + \tau_0^\alpha \frac{d^\alpha}{dt^\alpha}\right] V + \frac{1}{\sigma_\epsilon} V_S, \tag{2.8}$$

the fractional derivative arising out of the effects of viscous drag [44].

If the fractional derivative in (2.8) were absent we would obtain the classical Stokes solution

$$V(t) = V_S + (V_0 - V_S)e^{-t/\sigma_\epsilon}, \tag{2.9}$$

where σ_ϵ is the characteristic time of the motion, and V_S the final value of the velocity. The solution to equation (2.8) exhibits the same convergence as the classical Stokes solution this time with power-law rate convergence instead of exponential convergence of the solution to equation (2.9).

By the Basset equation we will mean a fractional differential equation of the form

$$Dy + aD^{\alpha}y + by = f,\ ab \neq 0,\ 0 < \alpha < 1. \tag{2.10}$$

2.3.3 The Bagley-Torvik Equation

When modelling the motion of a rigid plate in a Newtonian fluid a fractional differential equation of the form

$$D^2y + aD^{1+\alpha}y + by = f,\ ab \neq 0,\ 0 < \alpha < 1, \tag{2.11}$$

occurs [4]. By the Bagley-Torvik equation we will mean any fractional differential equation of the form give by (2.11).

Chapter 3

Fractional Calculus: Modern Theory

3.1 Motivation

In this chapter we give enough of the abstract theory of the fractional calculus to understand the origin of its computational challenges. Our development largely follows that found in [53].

We define fractional differential operators as they are understood in contemporary mathematics. The important difference between the Caputo and Riemann-Liouville version of the fractional derivative is fully explained.

In this section Abel's equation supplies the inspiration for the definition of the fractional integral operator. We begin with a number of definitions and a reassessment of Abel's integral. This enables us to give the basic ideas underlying the fractional calculus.

There are two ways in which the fractional calculus can be derived: firstly there is the *fractional integration and integer order differentiation* approach which we use here, and secondly there is a *limit of the fractional difference* approach, preferred by some authors particularly mathematical physicists [30].

3.2 Fubini's Theorem and Special Functions

Fubini's theorem is frequently appealed to in the fractional calculus to justify changing the order of integration in a multiple integral. Several special functions frequently occur in the study of the fractional calculus either in the definition of fractional differential operators, in the evaluation of fractional integrals, or as the solution to fractional differential equations.

3.2.1 Fubini's Theorem

The following Theorem, known as Fubini's Theorem, is used frequently in the fractional calculus.

Theorem 3.1 *([53] 9)*
Let $\Omega_1 = [a, b]$, $\Omega_2 = [c, d]$, $-\infty \leq a < b \leq \infty$, $-\infty \leq c < d \leq \infty$, *and let* $f(x, y)$ *be a measurable function defined on* $\Omega_1 \times \Omega_2$. *If at least one of the integrals*

$$\int_{\Omega_1} dx \int_{\Omega_2} f(x,y)dy, \quad \int_{\Omega_2} dy \int_{\Omega_1} f(x,y)dx, \quad \int_{\Omega_1 \times \Omega_2} f(x,y)dxdy, \tag{3.1}$$

is absolutely convergent then they coincide.

The case of Fubini's Theorem known as Dirchlet's formula is particularly useful. In this case we have

$$\int_b^a dx \int_a^x f(x,y)dy = \int_a^b \int_y^b f(x,y)dx, \tag{3.2}$$

assuming one of the integrals is absolutely convergent.

3.2.2 Special Functions

The gamma function occurs in the definition of fractional differential operators and the beta function may occur in evaluating fractional integrals. The Mittag-Leffler function and its generalisations occur as the solutions of fractional differential equations.

Definition 3.1 *([53] 15)*

The Euler integral of the second kind

$$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx, \ \mathrm{Re}\, z > 0, \tag{3.3}$$

is called the gamma-function. *The gamma-function is extended to the half-plane* $\mathrm{Re} \leq 0$, $z \neq 0, -1, \ldots$, *by analytic continuation of the integral. The reduction formula*

$$\Gamma(z) = z\Gamma(z), \ \mathrm{Re}\, z > 0, \tag{3.4}$$

is obtained from (3.3) through integration by parts. From (3.4) we can deduce that

$$\Gamma(z) = \frac{\Gamma(z+n)}{z(z+1)\ldots(z+n-1)}, \ \mathrm{Re}\, z > -n, \tag{3.5}$$

where $n = 1, 2, \ldots$, *and* $z \neq 0, -1, -2, \ldots$.

Definition 3.2 *([53] 17)*

The Euler integral of the first kind

$$B(z,w) = \int_0^1 t^{z-1}(1-t)^{w-1}dt, \mathrm{Re}\, z > 0, \ \mathrm{Re}\, w > 0, \tag{3.6}$$

is called the beta-function. *The beta-function can also be written in terms of gamma-functions*

$$B(z,w) = \frac{\Gamma(z)\Gamma(w)}{\Gamma(z+w)}. \tag{3.7}$$

Definition 3.3 *([53] 21)*

The Mittag-Leffler function *is an entire function defined by the series*

$$E_\alpha(t) = \sum_{k=0}^{\infty} \frac{t^k}{\Gamma(\alpha k + 1)}, \ \alpha > 0. \tag{3.8}$$

In [2] the Laplace-transform of the Mittag-Leffler function $E_\alpha(t^\alpha)$ is shown to be given by

$$\int_0^\infty e^{-pt} E_\alpha(t^\alpha) dt = \frac{1}{p - p^{1-\alpha}}, \ \mathrm{Re}\, p > 1. \tag{3.9}$$

Definition 3.4 *([53] 21)*

The generalised Mittag-Leffler function *is given by*

$$E_{\alpha,\beta}(t) = \sum_{k=0}^{\infty} \frac{t^k}{\Gamma(\alpha k + \beta)}, \ \alpha > 0, \ \beta > 0. \tag{3.10}$$

We can regard the generalised Mittag-Leffler function as the fractional generalisation of the exponential function. It occurs as the solution to the autonomous fractional differential equation.

Definition 3.5 *([25] 219)*

$$E_{(a_1,\ldots,a_n),b}(t_1,\ldots,t_n) = \sum_{k=0}^{\infty} \sum_{\substack{l_1+\cdots+l_n \\ l_1\geq 0,\ldots,l_n\geq 0}} (k; l_1,\ldots,l_n) \frac{\prod_{i=1}^n t_i^{l_i}}{\Gamma(b + \sum_{i=1}^n a_i l_i)}. \tag{3.11}$$

Where

$$(k; l_1,\ldots,l_n) = \frac{k!}{\prod_{i=1}^n (!l_i)}, \tag{3.12}$$

are the multinomial coefficients.

The multivariate Mittag-Leffler function occurs in the solution to multiterm fractional differential equations.

3.3 Abel's Integral Equation

Definition 3.6 *([53] 29)*

In this section we define Abel's equation as

$$\frac{1}{\Gamma(\alpha)}\int_0^x \frac{\phi(t)}{(x-t)^{1-\alpha}} = f(x),\ x > 0, \tag{3.13}$$

where $0 < \alpha < 1$.

The factor $1/\Gamma(\alpha)$ is introduced so that fractional differential operators produce the same results, when $\alpha \in \mathbb{Z}$, as the classical operations of integration and differentiation.

To solve the equation (3.13) [53] now takes the following approach: let $a > -\infty$ and suppose the equation is considered on the finite interval $[a, b]$. Substitute s for t and t for x and multiply both sides of the equation by $(x-t)^{-\alpha}$ and then integrating we have

$$\int_a^x \frac{dt}{(x-t)^{\alpha}}\int_a^t \frac{\phi(s)}{(t-s)^{1-\alpha}}ds = \Gamma(\alpha)\int_a^x \frac{f(t)}{(x-t)^{\alpha}}dt. \tag{3.14}$$

Interchanging the order of integration in the left-hand side, by means of Dirchlet's formula, we arrive at

$$\int_a^x \phi(s)ds \int_a^x \frac{dt}{(x-t)^{\alpha}(1-s)^{1-\alpha}} = \Gamma(\alpha)\int_a^x \frac{f(t)}{(x-t)^{\alpha}}dt. \tag{3.15}$$

The inner integral can be evaluated with the change of variable $t = s+\tau(x-s)$ and an application of the Beta function thus:

$$\begin{aligned}\int_s^x (x-t)^{-\alpha}(t-s)^{\alpha-1}dt &= \int_0^1 \tau^{\alpha-1}(1-\tau)^{-\alpha}d\tau \\ &= B(\alpha, 1-\alpha) \\ &= \Gamma(\alpha)\Gamma(1-\alpha).\end{aligned} \tag{3.16}$$

Therefore

$$\int_a^x \phi(s)ds = \frac{1}{\Gamma(1-\alpha)} \int_a^x \frac{f(t)}{(x-t)^\alpha} dt. \tag{3.17}$$

Which after differentiation gives

$$\phi(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_a^x \frac{f(t)}{(x-t)^\alpha} dt. \tag{3.18}$$

We conclude that if Abel's equation has a solution it is necessarily given by the above and is unique.

3.3.1 Absolute Continuity

The concept of absolute continuity is fundamental to determining the classes of functions that have fractional integrals and derivatives.

Definition 3.7 *([53] 2)*

A real valued function f defined on $[a,b]$ is said to be absolutely continuous on $[a,b]$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for every n disjoint open subintervals (a_k, b_k) of $[a,b]$, $n = 1, 2, \ldots$

$$\sum_{k=1}^{n} (b_k - a_k) < \delta \rightarrow \sum_{k=1}^{n} |f(b_k - f(a_k)| < \epsilon. \tag{3.19}$$

The space of such functions is denoted $AC([a,b])$.

It is known [34] that the space of functions $AC(\Omega)$ coincides with the space of primitives of Lebesgue summable functions:

$$f(x) \in AC \Leftrightarrow f(x) = c + \int_a^x \phi(t)dt, \int_a^b |\phi(t)|dt < \infty. \tag{3.20}$$

Therefore absolutely continuous functions have a summable derivative $f'(x)$ a.e..

Definition 3.8 *([53] 3)*

We denote by $AC^n(\Omega)$, where $n = 1, 2, \ldots$ and Ω is an interval, the space of functions $f(x)$ which have continuous derivatives up to order $n-1$ on Ω with $f^{(n-1)} \in AC(\Omega)$.

3.3.2 The Solvability of the Abel Equation in $L^1[a,b]$

In this subsection we present a theorem which provides more information about the conditions we must place on f such that Abel's equation is solvable.

First we observe that if

$$f_{1-\alpha}(x) = \frac{1}{\Gamma(1-\alpha)} \int_a^x \frac{f(t)}{(x-t)^\alpha} dt,$$

then if we integrate over $[a,b]$ and change the order of integration we obtain

$$\int_a^b |f_{1-\alpha}(x)| dx \leq \frac{1}{\Gamma(2-\alpha)} \int_a^b |f(t)|(b-t)^{1-\alpha} dt.$$

Therefore $f(x) \in L^1(a,b)$ implies $f_{1-\alpha}(x) \in L^1(a,b)$ as well.

Theorem 3.2 *([53] 31)*

Abel's integral equation for $\alpha \in [0,1]$ is solvable in $L^1(a,b)$ if and only if

$$f_{1-\alpha} \in AC([a,b]) \text{ and } f_{1-\alpha}(a) = 0. \tag{3.21}$$

Absolute continuity is an essential assumption since it is known that there are non-constant continuous functions which have zero derivative a.e..

Lemma 3.1 *([53] 32)*

If $f(x) \in AC([a,b])$, then $f_{1-\alpha} \in AC([a,b])$ and

$$f_{1-\alpha} = \frac{1}{\Gamma(2-\alpha)} \left[f(a)(x-a)^{1-\alpha} + \int_a^x f'(t)(x-t)^{1-\alpha} dt \right].$$

The next Corollary gives the form of the solution to Abel's integral equation.

Corollary 3.1 *([53] 32)*

If $f(x) \in AC([a,b])$, *then Abel's equation with* $0 < \alpha < 1$ *is solvable in* $L^1(a,b)$ *and its solution may be represented in the form*

$$\phi(x) = \frac{1}{\Gamma(1-\alpha)}\left[\frac{f(a)}{(x-a)^\alpha} + \int_a^x \frac{f'(s)}{(x-s)^\alpha}ds\right].$$

3.4 Fractional Integrals and Derivatives

An n-fold integral can be written as

$$\int_a^x dx \int_a^x dx \ldots \int_a^x \phi(t)dt = \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1}\phi(t)dt. \tag{3.22}$$

This formulation is generalised in the definition of the fractional integral.

Definition 3.9 *([53] 33)*

Let $\phi \in L^1(a,b)$. *The integrals*

$$(I_{a+}^\alpha \phi) = \frac{1}{\Gamma(\alpha)}\int_a^x \frac{\phi(t)}{(x-t)^{1-\alpha}}dt,\ x > a, \tag{3.23}$$

and

$$(I_{b-}^\alpha \phi) = \frac{1}{\Gamma(\alpha)}\int_x^b \frac{\phi(t)}{(x-t)^{1-\alpha}}dt,\ x < b, \tag{3.24}$$

where $\alpha > 0$, *are called fractional integrals of order* α. *These integrals are known as the* Riemann-Liouville *fractional integrals.*

The next theorem shows that fractional integral operators form a semigroup.

Theorem 3.3 *([53] 34)*

Fractional integration has the property

$$I_{a+}^{\alpha} I_{a+}^{\beta} \phi = I_{a+}^{\alpha+\beta} \phi, \; I_{b-}^{\alpha} I_{b-}^{\beta} \phi = I_{b-}^{\alpha+\beta} \phi, \; \alpha > 0, \beta > 0, \tag{3.25}$$

for all points for $\phi(t) \in C([a,b]$ *and if* $\phi \in L^1(a,b)$ *then for all points if* $\alpha + \beta \geq 1$ *and almost every point otherwise.*

We now give two definitions of the fractional derivative: firstly the Riemann-Liouville fractional derivative, and secondly the Caputo fractional derivative.

Definition 3.10 *([53] 33)*

For functions $f(x)$ *given in the interval* $[a,b]$ *with* $0 < \alpha < 1$, *the* left-handed Riemann-Liouville fractional derivative, *of order* α, *is given by*

$$(D_{a+}^{\alpha} f)(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_a^x \frac{f(t)}{(t-x)^{\alpha}} dt,$$

and the right-handed Riemann-Liouville fractional derivative, *of order* α, *is given by*

$$(D_{b-}^{\alpha} f)(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_x^b \frac{f(t)}{(t-x)^{\alpha}} dt.$$

In the Caputo fractional derivative the order of the fractional integration and differentiation is reversed

Definition 3.11 *([11] 530)*

For functions $f(x)$ *given in the interval* $[a,b]$ *with* $0 < \alpha < 1$, *the* left-handed Caputo fractional derivative, *of order* α, *is given by*

$$(D_{a+}^{\alpha} f)(x) = \frac{1}{\Gamma(1-\alpha)} \int_a^x \frac{f'(t)}{(t-x)^{\alpha}} dt,$$

and the right-handed Caputo fractional derivative, *of order* α, *is given by*

$$(D_{b-}^{\alpha} f)(x) = \frac{1}{\Gamma(1-\alpha)} \int_x^b \frac{f'(t)}{(t-x)^{\alpha}} dt.$$

We will be in approximating the solutions of fractional differential equations defined on an interval $[0, t]$ therefore we will be using the left-handed versions of the Riemann-Liouville and the Caputo fractional derivative with $a = 0$.

The next theorem shows how the two derivatives are related.

Theorem 3.4 *([25] 220)*

Let $f \in L^1(0, \infty)$ *and be* m*-time continuously differentiable for some* $m \in \mathbb{N}$ *and let* $m - 1 < \mu < m$. *Then the Riemann-Liouville and the Caputo fractional derivatives are connected by the relation:*

$$D^{\mu} f(t) = D^{\mu} f(t) + \sum_{k=0}^{m-1} \frac{f^{(k)}(0+)}{\Gamma(1+k-\mu)} t^{k-\mu}, \; t > 0. \tag{3.26}$$

Where the fractional derivative on the left hand side of equation (3.26) is understood to be a Riemann-Liouville fractional derivative and the fractional derivative on the right hand side is understood to be a Caputo fractional derivative.

Historical Note

The Riemann-Liouville fractional derivative has historical precedent over the Caputo fractional derivative and it is the fractional derivative used in the first definitions of fractional differential equations and in the attendant existence and uniqueness theory. Unless otherwise stated fractional derivatives in this chapter should be understood in the Riemann-Liouville sense. However the

use of the Riemann-Liouville fractional derivative limits the practical application of fractional differential equations to models with either homogenous initial conditions, i.e. $f^{(0)} = 0, \ldots, f^{(m)} = 0$, to eliminate singularities at $t = 0$, or to initial conditions specified in terms of fractional integrals, which lack a physical interpretation. Clearly this is an undesirable feature from the perspective of the mathematical modeller.

Following the spirit of Hadamard [26] *'The question, as set by most applications, does not consist in finding any solution of the differential equation, but in choosing, amongst all those possible solutions, a particular one defined by properly given accessory conditions.'* we are led to consider the ambiguity in the expression of fractional differential equations,which has no consequence in the classical case but is of the greatest moment in the fractional theory, arrising from the alternative sets of starting conditions.

The Caputo fractional derivative is better suited to the purposes of the mathematical modeller and therefore in subsequent chapters we will develop our numerical methods for approximating the solutions of fractional differential equations exclusively for fractional differential equations based on the Caputo fractional derivative.

The next lemma establishes a useful property of the Caputo fractional derivative at $t = 0$.

Lemma 3.2 *([21] 102)*

Let $y \in C^k[0,t]$ for some $k \in \mathbb{N}$, and let $\alpha \notin \mathbb{N}$ such that $0 < \alpha < k$. Then for the Caputo fractional derivative we have

$$D^{\alpha} y(t)|_{t=0} = 0. \tag{3.27}$$

The next lemma gives us more information on the difference between the Riemann-Liouville fractional derivative and the Caputo fractional derivative.

Lemma 3.3 *([53] 35)*

Let $f(x) \in AC([a,b])$, *then* $D^{\alpha}_{a+}f$ *and* $D^{\alpha}_{b-}f$ *exist a.e. for* $0 < \alpha < 1$. *Moreover* $D^{\alpha}_{a+}f$, $D^{\alpha}_{b-}f \in L_r(a,b)$, $1 \leq r < 1/\alpha$ *and we have*

$$D^{\alpha}_{a+}f = \frac{1}{\Gamma(1-\alpha)}\left[\frac{f(a)}{(x-a)^{\alpha}} + \int_a^x \frac{f'(t)}{(x-t)^{\alpha}}dt\right], \tag{3.28}$$

$$D^{\alpha}_{b-}f = \frac{1}{\Gamma(1-\alpha)}\left[\frac{f(b)}{(b-x)^{\alpha}} + \int_x^b \frac{f'(t)}{(x-t)^{\alpha}}dt\right]. \tag{3.29}$$

We see that for the Riemann-Liouville fractional derivative the function $(x-a)^{\alpha-1}$ plays the same role as the constant function does for the ordinary derivative and for the Caputo fractional derivative.

This is because both the ordinary derivative and the Caputo derivative, irrespective of its order, map constant functions on to 0. The Riemann-Liouville derivative of order α maps functions of the form $cx^{\alpha-1}$ on to 0.

The next set of definitions extends the range of α, the order of the fractional derivative, to all $\alpha > 0$.

Definition 3.12 *([53] 37)*

Let $\alpha = \alpha - [\alpha]$, *then for* $\alpha > 1$, *the* Riemann-Liouville fractional derivative *is defined by*

$$D^{\alpha}_{a+}f = (\frac{d}{dx})^{[\alpha]}D^{\{\alpha\}}_{a+}f = (\frac{d}{dx})^{[\alpha]}I^{1-\{\alpha\}}_{a+}f, \tag{3.30}$$

and

$$D^{\alpha}_{a+}f = (\frac{d}{dx})^{[\alpha]}D^{\{\alpha\}}_{a+}f = (\frac{d}{dx})^{[\alpha]}I^{1-\{\alpha\}}_{a+}f, \tag{3.31}$$

So, equivalently

$$D^{\alpha}_{a+}f = \frac{1}{\Gamma(n-\alpha)}(\frac{d}{dn})^n \int_a^x \frac{f(t)}{(x-t)^{\alpha-n+1}}dt,\ n = [\alpha]+1, \tag{3.32}$$

$$D^{\alpha}_{b-}f = \frac{1}{\Gamma(n-\alpha)}(\frac{d}{dn})^n \int_x^b \frac{f(t)}{(x-t)^{\alpha-n+1}}dt,\ n = [\alpha]+1. \tag{3.33}$$

For the Caputo fractional derivative we have

$$D^{\alpha}_{a+}f = \frac{1}{\Gamma(n-\alpha)} \int_a^x \frac{f^{(n)}(t)}{(x-t)^{\alpha}}dt,\ n = [\alpha]+1, \tag{3.34}$$

$$D^{\alpha}_{b-}f = \frac{1}{\Gamma(n-\alpha)} \int_x^b \frac{f^{(n)}(t)}{(x-t)^{\alpha}}dt,\ n = [\alpha]+1. \tag{3.35}$$

The next lemma gives the form an absolutely continuous function must take for a given level of differentiability.

Lemma 3.4 *([53] 39)*

The space $AC^n([a,b])$ consists of those and only those functions $f(x)$, which are represented in the form

$$f(x) = \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1}\phi(t)dt + \sum_{k=0}^{n-1} c_k(x-a)^k \tag{3.36}$$

where $\phi(t) \in L^1(a,b)$, c_k being arbitrary constants.

This follows from the definition of the space $AC^n([a,b])$ and from (3.20) and (3.22).

Theorem 3.5 *([53] 39)*

Let $\operatorname{Re}\alpha \geq 0$ and $f(x) \in AC^n([a,b])$, $n = [\operatorname{Re}\alpha]+1$. Then $D^{\alpha}_{a+}f$ exists a.e. and may be represented in the form

$$D^{\alpha}_{a+}f = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{\Gamma(1+k-\alpha)}(x-a)^{k-\alpha} + \frac{1}{\Gamma(n-\alpha)}\int_a^x \frac{f^{(n)}(t)}{(x-t)^{\alpha-n+1}}dt \tag{3.37}$$

Lemma 3.5 *([53] 40)*

Let $\phi(t) \in L^1(a,b)$. The homogeneous Abel integral equation $I_{a+}^{\alpha}\phi = 0$ has only the trivial solution $\phi(x) \equiv 0$ a.e. for all α, $\mathrm{Re}\,\alpha > 0$.

3.5 Fractional Integration and Differentiation as Reciprocal Operators

As we have already seen the Riemann-Liouville fractional derivative and the Caputo fractional derivative differ simply in the order in which fractional integration and ordinary differentiation occur but that this results in a significant difference in their properties. In this section we explore more fully the combination properties of fractional differential operators.

Definition 3.13 *([53] 43)*

Let $I_{a+}^{\alpha}(L^p)$, $\mathrm{Re}\,\alpha > 0$, denote the space of functions $f(x)$, represented by the left-side fractional integral of order α of a summable function: $f = I_{a+}^{\alpha}\phi$, $\phi \in L^p$, $1 \leq p < \infty$.

Theorem 3.6 *([53] 43)*

For $f(x) \in I_{a+}^{\alpha}(L^1)$, $\mathrm{Re}\,\alpha > 0$, it is necessary and sufficient that

$$f_{n-\alpha}(x) = I_{a+}^{n-\alpha} f \in AC^n([a,b]), \tag{3.38}$$

where $n = \mathrm{Re}\,\alpha + 1$ and that

$$f_{n-\alpha}^{(k)}(a) = 0,\ k = 0, 1, 2, \ldots, n-1. \tag{3.39}$$

Definition 3.14 *([53] 44)*

Let $\mathrm{Re}\,\alpha > 0$. A function $f(x) \in L^1(a,b)$ is said to have a summable fractional derivative D_{a+}^{α}, if $I_{a+}^{n-\alpha} f \in AC^n([a,b])$, $n = [\mathrm{Re}\,\alpha] + 1$.

In the next theorem we determine the consequences of altering the order of fractional differential operators.

Theorem 3.7 *([53] 44)*

Let $\operatorname{Re}\alpha > 0$. *Then the equality*

$$D^{\alpha}_{a+} I^{\alpha}_{a+} \phi = \phi(x) \tag{3.40}$$

is valid for any summable function $\phi(x)$ *while*

$$I^{\alpha}_{a+} D^{\alpha}_{a+} f = f(x) \tag{3.41}$$

is satisfied for

$$f(x) \in I^{\alpha}_{a+}(L^1). \tag{3.42}$$

If we assume that instead of (3.42) a function $f(x) \in L^1(a,b)$ *has a summable derivative* D^{α}_{a+}, *in the sense of definition 3.14, then (3.41) is not true in general and must be replaced by the result*

$$I^{\alpha}_{a+} D^{\alpha}_{a+} f = f(x) - \sum_{k=0}^{n-1} \frac{(x-a)^{\alpha-k-1}}{\Gamma(\alpha-k)} f^{(n-k-1)}_{n-\alpha}(a), \tag{3.43}$$

where $n = [\operatorname{Re}\alpha] + 1$ *and* $f_{n-\alpha}(x) = I^{n-\alpha}_{a+} f$. *In particular we have*

$$I^{\alpha}_{a+} D^{\alpha}_{a+} f = f(x) - \frac{f_{1-\alpha}(a)}{\Gamma(\alpha)} (x-a)^{\alpha-1}, \tag{3.44}$$

for $0 < \operatorname{Re}\alpha < 1$.

Corollary 3.2 *([53] 46)*

The following analogue of Taylor's Theorem

$$f(x) = \sum_{j=-n}^{n-1} \frac{(D^{\alpha+j}_{a+} f)(a)}{\Gamma(\alpha+j+1)} (x-a)^{\alpha+j} + R_n(x), \ \operatorname{Re}\alpha > 0, \tag{3.45}$$

is valid, where $R_n(x) = (I^{\alpha+n}_{a+} D^{\alpha+n}_{a+} f)(x)$ *and* $f(x)$ *is assumed to have a summable derivative* $D^{\alpha+a}_{a+} f$ *in the sense of definition 3.14.*

Equation (3.45) is a reformulation of the property given in equation (3.43).

Corollary 3.3 *([53] 46)*

The formula

$$\int_a^b f(x)(D_{a+}^{\alpha}g)(x)dx = \int_a^b g(x)(D_{b-}^{\alpha}f)(x)dx,\ 0 < \operatorname{Re}\alpha < 1, \qquad (3.46)$$

is valid under the assumption that $f(x) \in I_{b-}^{\alpha}(L^p), g(x) \in I_{a+}^{\alpha}(L^q), p^{-1}+q^{-1} \leq 1+\alpha$.

3.6 The Laplace Transform and Fractional Differential Operators

The Laplace transform is a useful tool for solving some classes of integro-differential equations. The Laplace transform has a particularly useful property with respect to convolution integrals. The fractional differential operators we have defined are examples of convolution integrals. Indeed as we will see later one method for obtaining an approximate solution to a fractional differential equation is derived by means of Laplace transform theory.

We review some well known but important properties of the Laplace transform here.

Definition 3.15 *([53] 27)*

Let f *be a function defined on the interval* $(0,\infty)$ *if the integral*

$$\mathcal{L}\{f\}(s) = \int_0^{\infty} f(t)e^{-st}dt$$

exists for some $s \in \mathbb{C}$ *where* $\operatorname{Re} s > 0$ *then* $\mathcal{L}\{f\}(s) = \bar{f}(s)$ *is called the* Laplace transform *of* f.

Theorem 3.8 *([53] 27)*

Let f be a function defined on $[0,\infty)$ then if its Laplace transform $\bar{f}$ exists then

$$f(x) = \mathcal{L}^{-1}\{\bar{f}(s)\}(x) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{sx}\bar{f}(s)ds,\ \gamma = \operatorname{Re} s > s_0. \tag{3.47}$$

As an example of the Laplace transform in action the Laplace transform of a power function is

$$\mathcal{L}\{x^n\}(s) = \frac{n!}{s^{n+1}}.$$

Theorem 3.9 *([53] 28)*

Let $f \in C^n[0,\infty)$ then the Laplace transform of $f^{(n)}$ is given by

$$\mathcal{L}\{f^{(n)}\}(s) = s^n\mathcal{L}\{f\} - \sum_{k=0}^{n-1} p^{n-k-1}f^k(0). \tag{3.48}$$

Let the convolution integral be represented by

$$f * g = \int_0^x f(x-t)g(t)dt.$$

Theorem 3.10 *([53] 28)*

The Laplace transform has the following convolution property

$$\mathcal{L}\{f * g\}(s) = \mathcal{L}\{f\}(s)\mathcal{L}\{g\}(s) = \bar{f}\bar{g}. \tag{3.49}$$

We can write the fractional integral as a convolution of the form

$$I_{0+}^{\alpha}f(x) = \frac{1}{\Gamma(\alpha)}\int_0^x \frac{f(t)}{(x-t)^{1-\alpha}}dt = \frac{1}{\Gamma(\alpha)}(\frac{1}{x^{1-\alpha}} * f(x)),$$

where $\operatorname{Re}\alpha > 0$. Thus we can write the Laplace transform of the fractional integral $I_{0+}^{\alpha}f$ as

$$\mathcal{L}\{I_0^{\alpha}f\}(s) = \frac{1}{s^{\alpha}}\mathcal{L}\{f\}(s) = \frac{\bar{f}}{s^{\alpha}}.$$

This enables us to write the fractional integral in terms of the Laplace transform operator and its inverse

$$I_{0+}^{\alpha} f(x) = \mathcal{L}^{-1} s^{-\alpha} \mathcal{L} f(x). \tag{3.50}$$

The next theorem shows how the generalised Mittag-Leffler function, given in definition 3.3, occurs as the inverse Laplace transform of a simple algebraic expression.

Theorem 3.11 *([53] 21)*

$$\mathcal{L}\{x^{\beta-1} E_{\alpha,\beta}((\lambda x)^{\alpha})\} = \frac{s^{\alpha-\beta}}{s^{\alpha} - \lambda^{\alpha}}. \tag{3.51}$$

3.7 Fractional Differential Equations

In this section we use the Caputo fractional derivative exclusively.

Definition 3.16 *([53] 829)*

The most general form of a fractional differential equation is an initial value problem of the form

$$F(t, y(t), D^{\alpha_1} y(t), D^{\alpha_2} y(t), \ldots, D^{\alpha_n} y(t)) = g(t), \tag{3.52}$$

where $y(t)$ is an unknown function and $0 < \alpha_1 < \cdots < \alpha_m$ and $y^{(k)}(0) = y_0^{(k)}$ for $k = 0, \ldots, n-1$.

For the purpose of this thesis we restrict ourselves to the two subclasses, simple non-linear fractional differential equations, given in definition 3.17, and the linear multiterm fractional differential equations, given in definition 3.18, since both these classes have a well-developed theory.

Definition 3.17 *([20],1)*

By a simple non-linear fractional differential equation we mean an initial values problem of the form

$$D^{\alpha}y(t) = f(t, y(t)), \tag{3.53}$$

where $n-1 < \alpha \leq n$ *for some* $n \in \mathbb{N}$, *where* $y^{(k)}(0) = y_0^{(k)}$ *for* $k = 0, 1, \ldots, m-1$.

Definition 3.18 *([25] 220)*

Let $0 \leq \beta_0 < \cdots < \beta_r < \beta_{r+1} < \cdots < \beta_p$, *where for each* i, $m_i - 1 < \beta_i \leq m_i$ *with* m_i *a non-negative integer, and let* $c_s \in \mathbb{R}$ *and for convenience* $c_p = 1$. *By a multiterm fractional differential equation we mean an initial value problem of the form*

$$\sum_{s=0}^{p} c_s D^{\beta_s} y = f,\ y^{(i)}(0) = y_0^{(i)} \text{ for } i = 0, \ldots, m_p - 1 \tag{3.54}$$

where $f \in L^1[0, \infty)$ *and if* β_p *is not an integer* f *has a continuous derivative.*

The simplest form of fractional differential equation is the autonomous fractional differential equation

$$D^{\alpha}y = -\beta y + f,\ y(0) = y_0. \tag{3.55}$$

The next theorem gives the form of the solution to the initial value problem given in definition 3.18.

Theorem 3.12 *([25] 220)*

The initial value problem given in definition 3.18 has a unique solution in the space of functions $L^1[0, \infty) \cap C^m[0, \infty)$ *of the form*

$$y(t) = y_f(t) + \sum_{k=0}^{m_p - 1} y_0^{(k)} u_k(t),\ t \geq 0, \tag{3.56}$$

where the function y_f has the form

$$y_f(t) = \int_0^t s^{\beta_p - 1} E_{(.),\beta_p}(s) f(t-s) ds, \tag{3.57}$$

this being the solution to the initial value problem given in 3.18 for $y^{(0)} = \cdots = y^{(m_p-1)}(0) = 0$, *and the system of functions*

$$u_k(t) = \frac{t^k}{k!} - \sum_{i=l_k+1}^{n} c_i t^{k+\beta_p-\beta_i} E_{k+1+\beta_p-\beta_i}(t),\ k = 0, \ldots, m_p - 1, \tag{3.58}$$

fulfils the initial conditions $u_0^{(l)} = \delta_{kl}$ *for* $k = 0, \ldots, m_p - 1$ *and* $l = 0, \ldots, m_p - 1$.

The function

$$E_{(.),\alpha}(t) = E_{(\beta_p-\beta_1,\ldots,\beta_p-\beta_{p-1}),\alpha}(-c_1 t^{\beta_p-\beta_1}, \ldots, -c_{p-1} t^{\beta_p-\beta_{p-1}}),$$

is a particular case of the multivariate Mittag-Leffler function and the natural numbers l_k, $k = 0, \ldots, m_p - 1$ *are determined from the condition*

$$m_{l_k} \geq k + 1,$$
$$m_{l_k} + 1 \leq k.$$

In the case $m_i \leq k$ *for* $i = 0, \ldots, m_p - 1$, *we let* $l_k = 0$, *and if* $m_i \geq k+1$ *for* $i = 0, \ldots, m_p - 1$, *we let* $l_k = p - 1$.

The next theorem establishes conditions for the simple non-linear fractional differential equation to have a unique solution.

Theorem 3.13 *([20], 2)*

Assume that $\mathcal{D} = [0, T^*] \times [y^{(0)} - \tau, y^{(0)} + \tau]$ *for some* $T^* > 0$ *and some* $\tau > 0$, *and let the function* $f : \mathcal{D} \to \mathbb{R}$ *be continuous in both variables, and let* $T = \min\{T^*, (\tau\Gamma(\alpha+1)/||f||_\infty)^{\frac{1}{\alpha}}\}$. *Then there exists a function* $y : [0, T] \to \mathbb{R}$ *solving the initial value problem given in definition 3.17.*

The next theorem establishes conditions for the function y in Theorem 3.13 to be unique.

Theorem 3.14 *([20], 3)*

Let $\mathcal{D}$, T^* and τ be as before. Let the function $f : \mathcal{D} \to \mathbb{R}$ be bounded on $\mathcal{D}$, continuous in the first variable and satisfy a Lipschitz condition in the second variable, with Lipschitz constant $L > 0$,

$$|f(x, y) - f(x, z)| \leq L|y - z|$$

independent of x, y, and z.

3.7.1 Dependence on Parameters

The dependence of the analytic solution, to the simple non-linear equation, to small perturbations of the parameters α, the initial conditions $y_0^{(k)}$ for $i = 0, \ldots, m-1$, and the function f, in the L_∞ norm, is considered in [20].

We give one of the theorems from [20] to make the nature of this result clear. We choose to quote the theorem which deals with small perturbations in the order of the fractional derivative.

Theorem 3.15 *([20])*

Assuming the condition of theorem 3.14 and let $\delta > 0$ be such that $m - 1 < \alpha - \delta < \alpha < m$. Assume that y and z are the uniquely determined solutions of the initial value problems

$$D^\alpha y(t) = f(t, y(t)),\ y(0) = y_0^{(0)}, \ldots, y^{(m-1)}(0) = y_0^{(m-1)}, \tag{3.59}$$

and

$$D^\alpha z(t) = f(t, z(t)),\ z(0) = z_0^{(0)}, \ldots, z^{(m-1)}(0) = z_0^{(m-1)}, \tag{3.60}$$

respectively. Then we have the relation

$$||y - z||_\infty = \mathcal{O}(\delta), \tag{3.61}$$

over any compact interval where both y and z exist.

The theorems relating to small perturbations of the initial conditions or the function f have the same format and reach the same conclusion, i.e. that the perturbation has linear consequences.

From this we concluded that fractional differential equations have unique solutions which are analytically well behaved and that the development of methods for obtaining numerical approximations to their solutions should be a feasible task for the computational mathematician given that such methods exist for ordinary differential equations.

Chapter 4

Approximate Solution of ODE's

4.1 Motivation

In this chapter we present some well established theory on the approximation of the solution of ordinary differential equations, our exposition largely follows that given in [32] and [36].

We start by defining various kinds of stability that a numerical method for approximating the solution of an ordinary differential equation may possess. We consider what implications the *stiffness* of a system of differential equations may have for the choice of a particular numerical method.

We then consider two classes of methods used for approximating the solutions of ordinary differential equations, these being linear multistep methods (LMM) and Runge-Kutta methods (RK). Methods for approximating the solution of a fractional differential equation have been derived from both LMM methods and RK methods.

In each case only a subclass of the methods used for ordinary differential equations can be used to approximate the solution of fractional differential equations. This is because of the intrinsic stiffness of fractional differential

equations.

4.2 General Concepts

We are interested in approximating solutions to ordinary differential equations of the form

$$y' = f(t, y), \quad t \geq t_0, \quad y(t_0) = y_0. \tag{4.1}$$

We assume f satisfies a Lipschitz condition in y since this is sufficient to guarantee the existence of a unique solution y to (4.1).

The numerical methods for approximating the solution of ordinary differential equations that we consider can all be written in the form

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h\phi_f(y_{n+k}, \ldots, y_n, t_n, h). \tag{4.2}$$

We assume that the functional ϕ_f satisfies a Lipschitz condition in $\mathbf{y}$ and that if $f \equiv 0$ then $\phi_f \equiv 0$, [36].

4.2.1 Stability

A necessary requirement of any numerical method is that it should converge to the true solution in a *reasonable* way. In this section we give some definitions that a method must satisfy to be usable.

Definition 4.1 *([32] 5)*

A method is said to be convergent *if for every ODE of the form of equation (4.1) and every* $t > 0$

$$\lim_{h \to 0} \max_{n=0,1,\ldots,\frac{t}{h}} \| y_{n,h} - y(t_n) \| = 0 \, t_n = nh.$$

Definition 4.2 *([32] 7)*

A time-stepping method of the form

$$y_{n+1} = Y_n(f, h, y_0, y_1, \ldots, y_n), \quad n = 1, 2, 3, \ldots, \tag{4.3}$$

is said to be of order p *if*

$$y(t_{n+1}) - Y_n(f, h, y(t_0), y(t_1), \ldots, y(t_n)) = \mathcal{O}(h^{p+1}). \tag{4.4}$$

Definition 4.3 *([36] 27)*

A method is said to be consistent *if, for all initial value problems of the form (4.1), the residual R_{n+k} given by*

$$R_{n+k} = \sum_{j=0}^{k} \alpha_j y(t_{n+j}) - h\phi_f(y(t_{n+k}), \ldots, y(t_n), t_n, h) \tag{4.5}$$

satisfies

$$\lim_{\substack{h \to 0 \\ x = t_0 + nh}} \frac{1}{h} R_{n+k} = 0. \tag{4.6}$$

By a perturbation of the initial value problem (4.1) we mean $(\delta t, \delta y_0)$ such that

$$z' = f(t, z) + \delta t, \; z(t_0) = y_o + \delta y_0. \tag{4.7}$$

Definition 4.4 *([36] 31)*

Let $(\delta t, \delta y_0)$ and $(\delta t^*, \delta y_0^*)$ be any two perturbation of (4.1) and let z and z^* be the resulting solutions. If there exists $S > 0$ such that for all $t^* \in [t_0, t]$, and all $\epsilon > 0$

$$|z(t^*) - z^*(t^*)| < S\epsilon \tag{4.8}$$

whenever

$$|\delta t - \delta t^*| \leq \epsilon \quad \text{and} \quad |\delta - \delta^*| \leq \epsilon, \tag{4.9}$$

then the initial value problem (4.1) is said to be totally stable.

Definition 4.5 *([36] 32)*

Let $\{\delta_n, n = 0, 1, \ldots, N\}$ *and* $\{\delta_n^*, n = 0, 1, \ldots, N\}$ *be any two perturbations of (4.1), and let* $\{z_n, n = 0, 1, \ldots, N\}$ *and* $\{z*_n, n = 0, 1, \ldots, N\}$ *be the resulting perturbed solutions. Then if there exist constants* $S > 0$ *and* $h_0 > 0$ *such that for all* $h \in (0, h_0]$ *and for all* $\epsilon > 0$

$$|z_n - z_n^*| < S\epsilon, \ 0 \leq n \leq N \tag{4.10}$$

whenever

$$|\delta_n - \delta_n^*| \leq \epsilon, \ 0 \leq n \leq N \tag{4.11}$$

we say that the method (4.2) is zero-stable.

Definition 4.6 *([32] 24)*

A polynomial $p(t)$ *satisfies the* root condition *if all the roots of* p *lie in the closed unit disc in* $\mathbb{C}$ *and those on the boundary of the unit disc are simple.*

Definition 4.7 *([36] 13)*

A polynomial $p(t)$ *is a* Schur *polynomial iff all the roots of* p *lie in the closed unit disc in* $\mathbb{C}$.

Theorem 4.1 *([36] 35)*

A necessary and sufficient condition for the method given by (4.2) to be zero stable is that it satisfies the root condition.

Theorem 4.2 *([36] 36)*

A necessary and sufficient condition for the method given by (4.2) to be convergent is that it be both consistent and zero stable.

4.2.2 Stiffness

Stiffness is a property that affects systems of ordinary differential equations which can lead to a rapid divergence between the approximate solution and the true solution. It is likely to occur if the system of ordinary differential equations models processes with vastly different rates of evolution, [32]. A long discussion in [36]concludes that stiffness is difficult to give an exact definition to.

If we consider a process such as viscoelasticity, which fractional differential equations may be used to model, we can see that two different physical processes with different time scales are involved; elasticity being a *fast* process and viscosity being a *slow* process. The Mittag-Leffler functions, which are the fractional counterpart of the exponential function, exhibit an initial transient which decays faster than any exponential function together with a tail which decays slower than any exponential function, i.e. we have a situation which exhibits vastly different rates of time evolution. Therefore we can only expect numerical methods for ordinary differential equations which work well with stiff systems to be candidate methods for generalising to numerical methods for fractional differential equations.

4.3 Linear Multistep Methods

Definition 4.8 *([36] 45)*

An s-step linear multistep method (LMM), for approximating the solution of an ordinary differential equation, is a relationship of the form

$$\sum_{j=0}^{s} \alpha_j y_{n+j} = h \sum_{j=0}^{s} \beta_j f_{n+j}, \tag{4.12}$$

where $\alpha_s = 1$ *and* $\alpha_0^2 + \beta_0^2 \neq 0$. *When* $b_s = 0$ *the method is said to be explicit otherwise it is said to be implicit.*

Definition 4.9 *([36] 45)*

We define the first and second characteristic polynomials ρ *and* σ *of a LMM to be*

$$\rho(\omega) = \sum_{m=0}^{s} a_m \omega^m \quad \textit{and} \quad \sigma(\omega) = \sum_{m=0}^{s} b_m \omega^m. \tag{4.13}$$

4.3.1 A-stability of LMM

In treating the A-stability of a given LMM we consider how it performs on the vector valued test system

$$y' = Ay, \; y(0) = y_0, \tag{4.14}$$

where the eigenvalues of A λ_i, $i = 1, 2, \ldots, m$, satisfy the condition

$$\operatorname{Re} \lambda_i < 0$$

and for all i, $\lambda_i = \lambda_j \rightarrow i = j$.

Then for the general solution of equation (4.14) we have $|y(t)| \rightarrow 0$ as $t \rightarrow \infty$. For a linear multistep method we must certainly require that the approximate solution has the property that $y_n \rightarrow 0$ as $n \rightarrow \infty$.

If we apply the LMM

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f_{n+j} \tag{4.15}$$

to (4.14) we have for $\{y_n\}$ the difference system

$$\sum_{j=0}^{k} (\alpha_j I - h\beta_j A) y_{n+j} = 0. \tag{4.16}$$

Assume, without loss of generality, that A is diagonal, perhaps with conjugate complex elements, then we can rewrite the above as a set of m *independent* equations

$$\sum_{j=0}^{k}(\alpha_j - \beta_j\lambda_t)\,{}^{t}y_{n+j} = 0,\ t = 1, 2, \ldots, m, \tag{4.17}$$

where $y_n = ({}^{1}y_n, {}^{2}y_n, \ldots, {}^{m}y_n)$. The general solution of each of the difference equations takes the form

$${}^{t}y_n = \sum_{s=1}^{m} c_{ts}r_s^n,\ t = 1, 2, \ldots, m, \tag{4.18}$$

where the c_{ts} are arbitrary complex numbers and r_s, $s = 1, 2, \ldots, m$ are the distinct roots of the characteristic polynomials

$$\sum_{j=0}^{k}(\alpha_j - h\beta_j\lambda_t)r^j. \tag{4.19}$$

These polynomials can be written in terms of the characteristic polynomials σ and ρ of the LMM

$$\pi(r, \lambda_t h) = \rho(r) - h\lambda_t\sigma(r),\ t = 1, 2, \ldots, m. \tag{4.20}$$

The polynomial $\pi(r, \lambda_t h)$ is called the stability polynomial of the method. The LMM will have the required convergence if all the roots r_s, $s = 1, 2, \ldots, k$ satisfy $|r_s| < 1$.

Definition 4.10 *([36] 70)*

A LMM is absolutely stable for $\lambda_t h$ if all the roots of the stability polynomial satisfy $|r_s| < 1$, $s = 1, 2, \ldots, k$ and absolutely unstable for $\lambda_t h$ otherwise.

Definition 4.11 *([36] 70)*

A LMM is said to have a region of absolute stability $\mathcal{R}_A$, where $\mathcal{R}_A \subset \mathbb{C}$,

and for all $a \in \mathcal{R}_A \rightarrow \mathrm{Re}\, a < 0$, *if it is stable for all* $\lambda_t h \in \mathcal{R}_A$. *The segment of the negative real axis given by* $\mathcal{R}_A \cap \mathbb{R}$ *is called the interval of absolute stability.*

4.3.2 Adam's Methods

We now consider a class of methods called Adams type methods. By an Adams type method we mean a LMM where

$$\rho(\zeta) = \zeta^k - \zeta^{k-1}. \tag{4.21}$$

When an Adams type method has the maximum possible accuracy it is known as an Adams-Bashforth method if it is explicit and an Adams-Moulton method if it is implicit.

To derive Adams type methods we integrate (4.1) over the interval $[t_{n+s-1}, t_{n+s}]$. Then by the fundamental theorem of calculus we have

$$y(t_{n+s}) = y(t_{n+s-1}) + \int_{t_{n+s-1}}^{t_{n+s}} y'(t)dt = y(t_{n+s-1}) + \int_{t_{n+s-1}}^{t_{n+s}} f(t, y(t))dt. \tag{4.22}$$

We now approximate the integral by interpolating for f with a polynomial approximation based on previously calculated values of the solution. Let

$$p(t) = \sum_{m=0}^{s-1} p_m(t) f(t_{n+m}, y_{n+m}), \tag{4.23}$$

where the *Lagrange interpolation polynomials*

$$p_m(t) = \prod_{\substack{l=0 \\ l \neq m}}^{s-1} \frac{t - t_{n+l}}{t_{n+m} - t_{n+l}} = \frac{(-1)^{s-1-m}}{m!(s-1-m)!} \prod_{\substack{l=0 \\ l \neq m}}^{s-1} \left(\frac{t - t_n}{h} - l\right), \tag{4.24}$$

are chosen such that

$$\begin{aligned} p_m(t) &= 1 \quad for \quad t = t_{n+m} \\ p_m(t) &= 0 \quad for \quad t \in \{t_n, \ldots, t_{n+s-1}\} - \{t_{n+m}\}. \end{aligned} \tag{4.25}$$

Clearly p is the unique polynomial of degree s which has this property because of the number of constraints.

Theorem 4.3 *([32] 362)*

Suppose that $f_l = f(\zeta_l)$, $l = 0, 1, \ldots, \nu$, *where* f *is a* $\nu+1$ *times differentiable function. Let* $a = \min_{0,1,\ldots,\nu} \zeta_i$ *and* $b = \max_{0,1,\ldots,\nu} \zeta_i$. *Then for every* $x \in [a, b]$ *there exists* $\eta = \eta(x) \in [a, b]$ *such that*

$$p(x) - f(x) = \frac{1}{(\nu+1)!} f^{\nu+1}(\eta) \prod_{k=0}^{\nu} (x - \zeta_k). \tag{4.26}$$

By Theorem 4.3 if f is $s+1$ times differentiable for h sufficiently small we have

$$p(t) = y'(t) + \mathcal{O}(h^s), \quad t \in [t_{n+m-1}, t_{n+m}].$$

Substituting (4.23) in (4.3), replacing $y(t_{n+s-1})$ by y_{n+s-1}, and integrating along an interval of length h incurs an error of $\mathcal{O}(h^{s+1})$. Therefore the method

$$y_{n+s} = y_{n+s-1} + h \sum_{m=0}^{s-1} b_m f(t_{n+m}, y_{t+m}) \tag{4.27}$$

where

$$b_m = h^{-1} \int_{t_{n+s-1}}^{t_{n+s}} p_m(t) dt = h^{-1} \int_0^h p_m(t + t_{n+s-1}) dt, \quad m = 0, 1, \ldots, \tag{4.28}$$

is of order s. For $s = 1$ this reduces to the Euler method.

4.3.3 Order and Convergence of LMM

Let the general s-step method be written in the form

$$\sum_{m=0}^{s} a_m y_{n+m} = h \sum_{m=0}^{s} b_m f(t_{n+m}, y_{n+m}), \quad n = 0, 1, \ldots, \tag{4.29}$$

where $a_m, b_m, m = 0, 1, \ldots, s$, are given constants independent of h, n and the underlying differential equation. We will take $a_s = 1$.

Theorem 4.4 *([32] 22)*

Let y be analytic and its radius of convergence exceed sh then the multistep method (4.29) is of order $p \geq 1$ if and only if there exists $c \geq 0$ such that

$$\rho(\omega) - \sigma(\omega) \ln \omega = c(\omega - 1)^{p+1} + \mathcal{O}(|\omega - 1|^{p+2}),\ \omega \to 1. \tag{4.30}$$

Theorem 4.5 ***Dahlquist's equivalence theorem*** *([32] 24)*

Suppose that the error in the starting values $y_1, y_2, \ldots, y_{s-1}$ tends to zero as $h \to 0+$. The multistep method (4.30) is convergent iff it is of order $p \geq 1$ and the polynomial ρ is Schur.

The next theorem is known as *The First Dahlquist Barrier.* It establishes the maximum order possible for a LMM which also possesses zero-stability.

Theorem 4.6 *([36] 55)*

No zero-stable linear k-step method can have order exceeding $k+1$ when k is odd and $k+2$ when k is even.

To further describe the properties of LMM we need some more definitions concerning stability.

Definition 4.12 *([36] 224)*

A method is said to be A-stable if $\mathcal{R}_A \supseteq \{\lambda_t h | \operatorname{Re} \lambda_t h < 0\}$.

Definition 4.13 *([36] 225)*

A method is said to be $A(\alpha)$-stable if for some $\alpha(0, \pi/2)$ we have $\mathcal{R}_A \supseteq \{\lambda_t h | -\alpha < \pi - \lambda_t h < \alpha\}$.

Definition 4.14 *([36] 225)*

A method is said to be A_0-stable if $\mathcal{R}_A \supseteq \{\lambda_t h | \operatorname{Re} \lambda_t h < 0 \ \operatorname{Im} \lambda_t h = 0\}$.

The next theorem is known as *The Second Dahlquist Barrier.* It establishes the maximum order possible for a LMM which also possesses absolute stability.

Theorem 4.7 *([36] 233)*

(i) An explicit LMM cannot be A-stable.

(ii) The order of an A-stable LMM cannot exceed 2.

(iii) The second order A-stable linear multistep method with smallest error constant is the Trapezoidal Rule.

4.3.4 Backward Differentiation Formulae

The backward differentiation formulae (BDF) are the best known LMM that can be developed into fractional linear multistep methods.

Definition 4.15 *([32] 27)*

An s-order, s-step method is said to be a BDF if $\sigma(\omega) = \beta\omega^s$ *for some* $\beta \in \mathbb{R} - \{0\}$.

Lemma 4.1 *([32] 27)*

For a BDF we have

$$\beta = (\sum_{m=1}^{s} \frac{1}{m})^{-1} \quad and \quad \rho(\omega) = \beta \sum_{m=1}^{s} \frac{1}{m} \omega^{s-m}(\omega - 1)^m. \tag{4.31}$$

4.4 Runge-Kutta Methods

We can rewrite the differential system

$$y' = f(t), \;\; t \geq t_o, \;\; y(t_0) = y_0 \tag{4.32}$$

k	α_6	α_5	α_4	α_3	α_2	α_1	α_0	β_k	p
1						1	-1	1	1
2					1	$-\frac{4}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	2
3				1	$-\frac{18}{11}$	$\frac{9}{11}$	$-\frac{2}{11}$	$\frac{6}{11}$	3
4			1	$-\frac{48}{25}$	$\frac{36}{25}$	$-\frac{16}{25}$	$\frac{3}{25}$	$\frac{12}{25}$	4
5		1	$-\frac{300}{137}$	$\frac{300}{137}$	$-\frac{200}{137}$	$\frac{75}{137}$	$-\frac{12}{137}$	$\frac{60}{137}$	5
6	1	$-\frac{360}{147}$	$\frac{450}{147}$	$-\frac{400}{147}$	$\frac{225}{147}$	$-\frac{72}{147}$	$\frac{10}{147}$	$\frac{60}{147}$	6

Table 4.1: Coefficients of the BDF

in integral form

$$y(t) = y_0 + \int_{t_0}^{t} f(t)dt, \tag{4.33}$$

and obtain an approximate solution by applying a quadrature method to estimate the integral on the right hand side of equation (4.33). In Runge-Kutta methods we apply this methodology to systems of the form

$$y' = f(t, y), \quad t \geq t_0, \quad y(t_0) = y_0. \tag{4.34}$$

A weight function is a nonnegative function ω defined on the interval (a, b), such that

$$0 < \int_a^b \omega(t)dt < \infty, \quad | \, 0 < \int_a^b t^j \omega(t)dt \, | < \infty, \quad j = 1, 2, \ldots. \tag{4.35}$$

We make the approximation

$$\int_a^b f(t)\omega(t)dt \approx \sum_{j=1}^{\nu} b_j f(c_j), \tag{4.36}$$

where b_i and c_i are independent of the function f and are called the quadrature weights and nodes respectively.

The next theorem is known as the Peano kernel theorem.

Theorem 4.8 *([50] 43)*

Let $f \in C^p[a,b]$ and let the linear functional $F(f)$ be approximated by the linear functional $G(f)$ such that the error

$$E(f) = F(f) - G(f) \tag{4.37}$$

vanishes for all polynomials of degree $p-1$ or less. Then

$$E(f) = \int_a^b f^{(p)}(t)K(t)dt \tag{4.38}$$

where

$$K(t) = \frac{1}{n!}E_x[(x-t)_+^n] \tag{4.39}$$

and

$$(x-t)_+^n = \begin{cases} (x-t)^n & x \geq t \\ 0 & x < t \end{cases} . \tag{4.40}$$

The function $K(t)$ is called called the Peano kernel *for the linear functional E.*

It can be shown, by use of the Peano kernel theorem, that if a quadrature method is exact for polynomials of degree $p-1$ then for all $f \in C^p[a,b]$

$$| \int_a^b f(t)\omega(t)dt - \sum_{j=1}^{\nu} b_j f(c_j) | \leq c \max_{a \leq t \leq b} | f^{(p)}(t) | \tag{4.41}$$

where c is a constant independent of f. Denoting the set of all real polynomials of degree m by $\mathbb{P}_{p-1}$ then (4.35) is of order p if it is exact for every $f \in \mathbb{P}_{p-1}$.

We give the proof of the following lemma since some of its equations are referred to latter in this section.

Lemma 4.2 *([32] 34)*

Given a set of distinct nodes $c_1, c_2, \ldots, c_\nu$, it is possible to find a unique set of weights $b_1, b_2, \ldots, b_\nu$, such that (4.36) is exact for some $p \geq \nu$.

Proof The set $T_{\nu-1} = \{1, t^1, t^2, \ldots, t^{\nu-1}\}$ is a linearly independent basis for the space $\mathbb{P}_{\nu-1}$ thus, since the quadrature formula is a linear operator, if it is exact for each element of $T_{\nu-1}$ it will be exact for all elements of $\mathbb{P}_{\nu-1}$. Integrating for each element of $T_{\nu-1}$ gives

$$\sum_{j=1}^{\nu} b_j c_j^m = \int_a^b t^m \omega(t) dr, \quad m = 0, 1, \ldots, \nu - 1, \tag{4.42}$$

which is a system of ν equations in the unknowns $b_1, b_2, \ldots, b_\nu$, which form a non-singular Vandermond matrix.

Using the nodes $c_1, c_2, \ldots, c_\nu$ as the interpolating points of Lagrange polynomials the weights $b_1, b_2, \ldots, b_\nu$ can be calculated exactly.

Let

$$p_j(t) = \prod_{\substack{k=1 \\ k \neq j}}^{\nu} \frac{t - c_k}{c_j - c_k}, \quad j = 1, 2, \ldots, \nu \tag{4.43}$$

then since for all $g(t) \in \mathbb{P}_{\nu-1}$

$$\sum_{j=1}^{\nu} p_j(t) g(c_j) = g(t) \tag{4.44}$$

we have

$$\sum_{j=1}^{\nu} \int_0^1 p_j(t) \omega(t) dt c_j^m = \int_0^1 \sum_{j=1}^{\nu} [p_j(t) c_j^m] \omega(t) dt = \int_0^1 t_j^m \omega(t) dt \tag{4.45}$$

for every $m = 0, 1, \ldots, \nu - 1$. Therefore

$$b_j = \int_0^1 p_j(t) \omega(t) dt, \quad j = 1, 2, ..., \nu. \tag{4.46}$$

4.4.1 Quadrature and Orthogonal Polynomials

Quadrature methods which use equally spaced nodes in the interval of integration are called Newton-Cotes methods. Higher order methods can be achieved by the use of orthogonal polynomials.

Let $L^2(\omega)$ be the set of functions such that

$$\int_a^b \mid f(t) \mid^2 \omega(t)dt < \infty \tag{4.47}$$

then ω defines an inner product

$$< f, g >_\omega = \int_a^b f(t)g(t)\omega(t)dt \tag{4.48}$$

on $L^2(\omega)$.

Definition 4.16 *([32] 35)*

Let $p_m \in (P)_m$, $p_m \neq 0$ then p_m is an m-th order polynomial with respect to ω if

$$< p_m, p >= 0, \quad \forall \ p \in \mathbb{P}_{m-1} \tag{4.49}$$

We call a polynomial monic if the coefficient of its highest power is equal to 1.

If we let $b = -a = 1$, and $\alpha, \beta > -1$, the sets of underlying orthogonal polynomials with respect to the weight function $\omega(t) = (1-t)^\alpha(1+t)^\beta$ are know as Jacobi polynomials.

Lemma 4.3 *([32] 36)*

All m zeros of an orthogonal polynomial p_m reside in the interval (a, b) and they are simple.

Theorem 4.9 *([32] 36)*

Let $c_1, c_2, \ldots, c_\nu$ be the zeros of p_ν and let $b_1, b_2, \ldots, b_\nu$ be the solutions of the Vandermond system (4.46). Then

(i) The quadrature method (4.36) is of order 2ν;

(ii) No other quadrature can exceed this order.

Methods which achieve the conditions of the above theorem are known as Gauss methods.

4.4.2 Explicit Runge-Kutta Methods

To apply a quadrature formula to the ODE (4.1) we begin by integrating between t_n and t_{n+1} to obtain

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t))dt = y(t_n) + h\int_0^1 f(t_n + ht, y(t_n + ht)dt.$$

If we replace the integral with a quadrature such as

$$y_{n+1} = y_n + h\sum_{j=1}^{\nu} b_j f(t_n + c_j h, y(t_n + c_j h)), \quad n = 0, 1, \ldots, \tag{4.51}$$

and if we knew y at the nodes $t_n + c_1, t_n + c_2, \ldots, t_n + c_\nu$ we would have a solution.

To effect a viable Runge-Kutta method we must approximate $y(t_n + c_j h)$ for $j = 1, 2, \ldots \nu$. Let $\xi_j \approx y(t_n + c_j h)$ for $j = 1, 2, \ldots \nu$. For an explicit

method we set $c_1 = 0$. Then we have

$$\begin{aligned}
\xi_1 &= y_n \\
\xi_2 &= y_n + h a_{2,1} f(t_n, \xi_1) \\
&\vdots \\
\xi_\nu &= y_n + h \sum_{i=1}^{\nu-1} a_{\nu,i} f(t_n + c_i h, \xi_i) \\
y_{n+1} &= y_n + h \sum_{j=1}^{\nu} b_j f(t_n + c_j h, \xi_j).
\end{aligned} \tag{4.52}$$

Let A be a $\nu \times \nu$ real matrix, called the RK matrix or *Butcher array*, with elements equal to $(a_{j,i})_{j,i=1,2,\ldots,\nu}$ as defined above and zero otherwise. And let

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_\nu \end{bmatrix} \text{ and } \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_\nu \end{bmatrix}. \tag{4.53}$$

The vector $\mathbf{b}$ is called the RK weights, and the vector $\mathbf{c}$ is called the RK nodes and we say that (4.52) has ν stages or alternatively is a ν stage method. RK-methods are commonly written as *RK-tableaux* which have the following form

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array}. \tag{4.54}$$

Examples of explicit two-stage RK-methods are

$$\begin{array}{c|cc} 0 & & \\ \frac{1}{2} & \frac{1}{2} & \\ \hline & 0 & 1 \end{array} \qquad \begin{array}{c|cc} 0 & & \\ \frac{2}{3} & \frac{2}{3} & \\ \hline & \frac{1}{4} & \frac{3}{4} \end{array} \qquad \begin{array}{c|cc} 0 & & \\ 1 & 1 & \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}. \tag{4.55}$$

The coefficients of some RK-methods, of up to order four, can be obtained by Taylor series expansions. For more general and higher order methods we

need to use techniques derived from graph theory to keep track of all the combinations of derivatives that occur.

The following condition must be fulfilled so that the explicit RK-method has order one

$$\sum_{i=1}^{j-1} a_{j,i} = c_j, \ j = 2, 3, \ldots, \nu, \tag{4.56}$$

otherwise we would not be able to approximate the solution of $y' \equiv 1$.

Explicit RK-methods of fourth-order, and below, require the same number of stages as the order of the method. Fifth order ERK-methods require six-stages. Higher order methods also suffer from stage inflation.

4.4.3 Implicit Runge-Kutta Methods

In explicit RK methods the vector functions ζ_j are only allowed to depended upon $\zeta_1, \ldots, \zeta_{j-1}$. In implicit Runge-Kutta methods the vector function ζ_j may depend upon all the ζ_i for $i = 1, \ldots, \nu$. Thus the scheme (4.52) is generalised to

$$\begin{aligned} \xi_j &= y_n + h \sum_{i=1}^{\nu} a_{j,i} f(t_n + c_i h, \xi_i), \\ y_{n+1} &= y_n + h \sum_{j=1}^{\nu} b_j f(t_n + c_j h, \xi_j), \end{aligned} \tag{4.57}$$

where $A = (a_{j,i})_{j,i=1,2,\ldots,\nu}$ is now an arbitrary square matrix.

The convention

$$\sum_{i=1}^{\nu} a_{j,i} = c_j, \ \ j = 1, 2, \ldots, \nu, \tag{4.58}$$

is required to ensure that the method is of nontrivial order. For a general RK matrix A (4.57) yields a system of νd coupled algebraic equations where $y \in \mathbb{R}^d$.

4.4.4 Realisation of Implicit RK Through Collocation

There exists a particular class of IRK-methods which are particularly suitable for generalisation to the numerical approximation of fractional differential equations. We describe these methods in this section.

Recollect that in explicit RK-methods the vector functions ζ_j are only allowed to depend upon $\zeta_1, \ldots, \zeta_{j-1}$. In the implicit RK methods we consider here the vector functions ζ_j which are only allowed to depend upon $\zeta_1, \ldots, \zeta_j$ for $j = 1, 2, \ldots, \nu$.

Suppose the solution to equation (4.1) has been calculated up to (t_n, y_n) and we wish to calculate (t_{n+1}, y_{n+1}) , where $t_{n+1} = t_n + h$. In a collocation method we choose collocation parameters $\{c_1, c_2, \ldots, c_\nu\} \subset [0, 1]$ such that $i < j \rightarrow c_i < c_j$ and calculate a ν-th degree polynomial such that

$$\begin{aligned} u(t_n) &= y_n, \\ u'(t + n + c_j h) &= f(t_n + c_j h, u(t_n + c_j h)), \quad j = 1, 2, \ldots, \nu \end{aligned} \tag{4.59}$$

and then make the approximation

$$y_{n+1} = u(t_{n+1}). \tag{4.60}$$

Such a scheme is equivalent to an implicit Runge-Kutta method.

Lemma 4.4 *([32] 43)*

Set

$$q(t) = \prod_{j=1}^{\nu}(t - c_j), \quad q_l(t) = \frac{q(t)}{t - c_l}, \quad l = 1, 2, \ldots, \nu, \tag{4.61}$$

and let

$$a_{j,i} = \int_0^{c_j} \frac{q_i(t)}{q_i(c_i)} dt, \quad j, i = 1, 2, \ldots, \nu, \tag{4.62}$$

$$b_{j,i} = \int_0^1 \frac{q_i(t)}{q_i(c_i)} dt, \quad j = 1, 2, \ldots, \nu. \tag{4.63}$$

Then the collocation method (4.59) is identical to the IRK-method

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array}. \tag{4.64}$$

Definition 4.17 *([32] 45)*

Suppose we have a smoothly differentiable function v as a candidate solution for (4.1) and that $v(t_0) = y_0$. Then the defect is given by

$$d(t, v) = v'(t) - f(t, v(t)). \tag{4.65}$$

Clearly $d(t, y) = 0$ thus we expect a small value of $\| d(t, v) \|$ to imply a small error when we interpret $v(t)$ as the solution to (4.1).

For a system of linear equations we have

$$y' = \Lambda y, \quad y(t_0) = y_0, \tag{4.66}$$

and for any candidate solution v the associated system

$$v' = \Lambda v + d(t), \quad t \geq t_0, v(t_0) = v_0. \tag{4.67}$$

These have as their exact solutions

$$\begin{aligned} y(t) &= e^{(t-t_0)\Lambda} y_0. \quad t \geq t_0, \\ v(t) &= e^{(t-t_0)\Lambda}(v_0 - y_0) + \int_{t_0}^{t} e^{(t-x)\Lambda} d(x) dx, \quad t > t_0. \end{aligned} \tag{4.68}$$

The difference in these gives

$$v(t) - y(t) = e^{(t-t_0)\Lambda}(v_0 - y_0) + \int_{t_0}^{t} e^{(t-x)\Lambda} d(x) dx, \quad t \geq t_0, \tag{4.69}$$

hence the error can be expressed entirely in terms of $v_0 - y_0$ and d.

Theorem 4.10 *(The Alekseev-Grobner lemma) ([32] 45)*

Let v be a smoothly differentiable function that obeys the initial condition $v(t_0) = y_0$. Then

$$v(t) - y(t) = \int_{t_0}^{t} \Phi(t, x, v(x)) d(x, v) dx, \quad t \geq t_0, \tag{4.70}$$

where Φ is the matrix of partial derivatives of the solution of the ODE $w' = f(t, w)$, $w(t) = v(t)$, with respect to $v(t)$.

Theorem 4.11 *([32] 46)*

Suppose that

$$\int_0^1 q(t) t^j dt = 0, \quad j = 1, 2, \ldots, m-1, \tag{4.71}$$

for some $m \in \{0, 1, \ldots, \nu - 1\}$ with q as in lemma 4.4. Then the collocation method (4.59) is of order $\nu + m$.

Corollary 4.1 *([32] 46)*

Let $c_1, c_2, \ldots, c_\nu$ be the zeros of the polynomial $\tilde{P}_\nu \in \mathbb{P}_\nu$ that is orthogonal with the weight function $\omega(t) \equiv 1$, $0 \leq t \leq 1$. Then the underlying collocation method (4.59) has order 2ν.

Chapter 5

Grünwald-Letnikov Fractional Derivative

5.1 Motivation

In this chapter we present the simplest method for approximating a fractional integral. This method is based on using the Grünwald-Letnikov version of the fractional derivative. We see that is insufficient to provided a method for approximating the solution to a fractional differential equation.

5.2 The Grünwald-Letnikov Fractional Derivative

Definition 5.1 *([18] 449)*

The Grünwald-Letnikov fractional derivative, based on backward differences, is given by

$$D^{\alpha}y(t)|_{t=b} = \lim_{N\to\infty}(\frac{b-a}{N})^{-\alpha}\sum_{k=0}^{N}\frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)}y(b-k\frac{(b-a)}{N}). \qquad (5.1)$$

this can be shown to be equivalent to the Riemann-Liouville fractional derivative [46].

The simplest method for approximating the fractional derivative is to replace the infinite sum in the Grünwald-Letnikov fractional derivative with a finite sum for some integer N. For the function $y(t)$ defined on the interval $[a, b]$ the approximate fractional derivative K^α at b is given by

$$K^\alpha y(t)|_{t=b} = (\frac{b-a}{N})^{-\alpha} \sum_{k=0}^{N} \frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)} y(b - k\frac{(b-a)}{N}). \tag{5.2}$$

Without loss of generality we let $a = 0$, $b = j/N$, and discretise $[0, 1]$ as $0, \frac{1}{N}, \cdots, \frac{j}{N}, \cdots, 1$. Putting $t = j/N$, in equation (5.2), we have

$$K^\alpha y(\frac{j}{N}) = (\frac{1}{N})^{-\alpha} \sum_{k=0}^{j} \frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)} y(\frac{j-k}{N}). \tag{5.3}$$

So writing $y(j/N)$ as y_j this is

$$K^\alpha y_j = (\frac{1}{N})^{-\alpha} \sum_{k=0}^{j} \frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)} y_{j-k}. \tag{5.4}$$

Theorem 5.1 *([18] 453)*

Let $f \in L^1[0, t]) \cup C^1[0, t]$ *and let* $h = 1/N$ *then as* $h \to 0$

$$(D^\alpha - K^\alpha) f(t) = \mathcal{O}(h). \tag{5.5}$$

In figures 5.1 and 5.2 the upper curve is the fractional derivative of t at the interpolation points and the lower curve is the approximate fractional derivative of t at the interpolation points. It can be seen that equation (5.1) produces an approximation of the fractional derivative which converges to the true value.

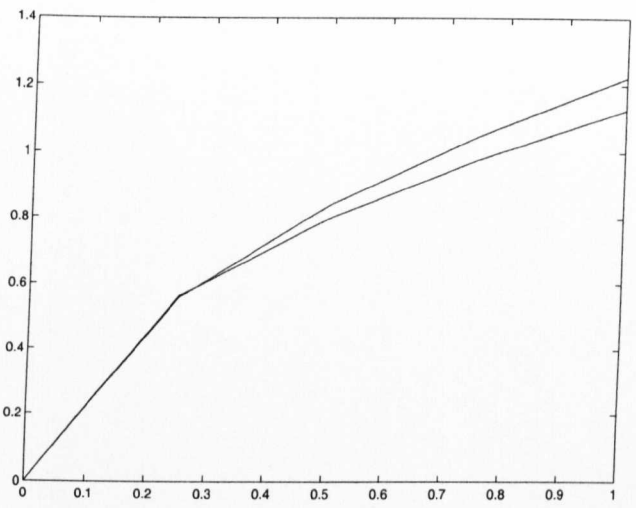

Figure 5.1: Approximate fractional derivative of $y = t$ for step length of $1/4$ where $\alpha = 0.5$.

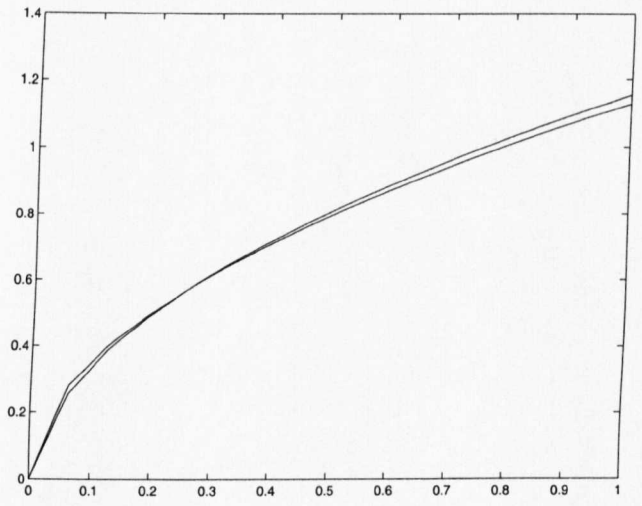

Figure 5.2: Approximate fractional derivative of $y = t$ for step length of $1/16$ where $\alpha = 0.5$.

5.2.1 The Autonomous Fractional Differential Equation

The above suggests that to approximate the linear autonomous fractional differential equation

$$D^{\alpha}y = -\lambda y,\ y(0) = y_0,\ \lambda > 0, \tag{5.6}$$

we might try

$$-\lambda y_j = (\frac{1}{N})^{-\alpha} \sum_{k=0}^{j} \frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)} y_{j-k}. \tag{5.7}$$

Solving for y_j gives

$$y_j = \frac{-N^{\alpha}}{(\lambda + N^{\alpha})} \sum_{k=1}^{j} \frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)} y_{j-k}. \tag{5.8}$$

For the forced system

$$D^{\alpha}y = -\lambda x + f,\ y(0) = y_0,\ \lambda > 0, \tag{5.9}$$

we have

$$y_j = \frac{1}{(\lambda + N^{\alpha})} (f_j - N^{\alpha} \sum_{k=1}^{j} \frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)} y_{j-k}). \tag{5.10}$$

Attempts to use (5.10) to approximate the system

$$D^{\alpha}y(t) = -y(t) + t + \frac{t^{0.5}}{\Gamma(1.5)},\ y(0) = 0, \tag{5.11}$$

are plotted in figures 5.3 and 5.4. As we can see, since (5.11) has as its solution $y(t) = t$ there is no evidence of the approximation converging to the straight line. Therefore we conclude that equation (5.10) does not produce a good approximation to the true solution of (5.11). This is because we require several of the initial terms in the sum in equation (5.1) if it is to be a good approximation of the fractional derivative.

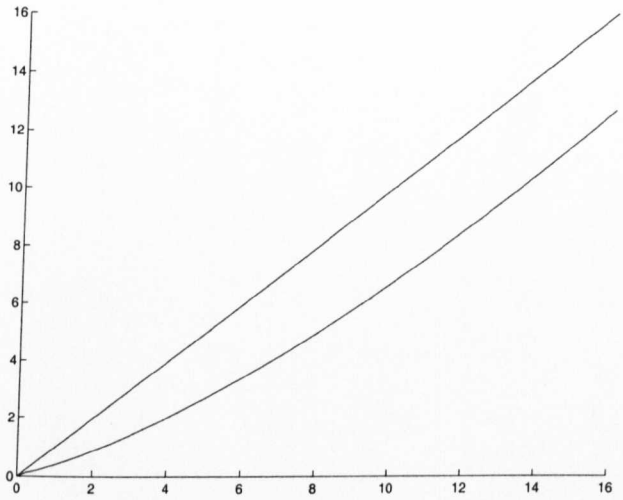

Figure 5.3: Approximate solution to 5.11 with a step length of 1/4 where $\alpha = 0.5$.

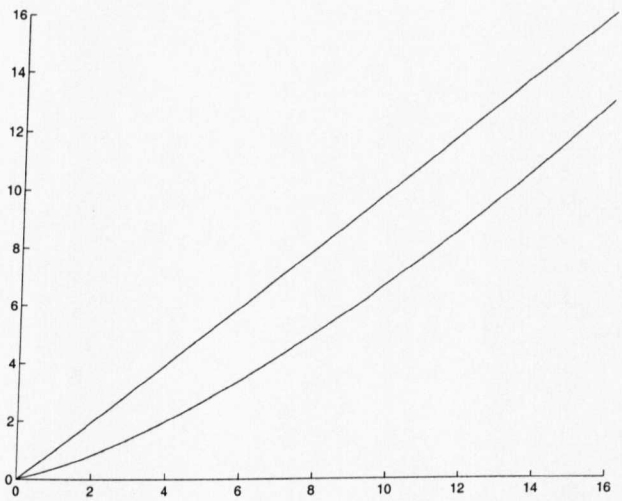

Figure 5.4: Approximate solution to 5.11 with a step length of 1/16 where $\alpha = 0.5$.

Chapter 6

Fractional Trapezium Rule

6.1 Motivation

In this chapter we give a full exposition of a method for numerically approximating the solution to a fractional differential equation which was first presented in [12]. By fully developing the method here we can easily extend it to fit the more sophisticated approach to mesh point distribution which we will present later.

Two principal techniques are used in this method, firstly the trapezium rule is extended to estimate the fractional integral by interpreting it as a product quadrature and secondly the singularity at the origin is dealt with by treating the fractional integral as a Hadamard finite part integral [18], [13]. For a step length of $h = 1/n$, $n \in \mathbb{N}$ the fractional trapezium rule has $\mathcal{O}(h^{2-\alpha})$ convergence and requires $\mathcal{O}((nt)^2)$ computational effort to calculate an approximate solution over an interval $[0, t]$.

Throughout this chapter we will use (6.1) as our test equation

$$D^{\alpha}y(t) = -\lambda y(t) + f(t),\ y(0) = y_0 \in \mathbb{R},\ \lambda > 0, \tag{6.1}$$

where $0 < \alpha < 1$.

6.2 Trapezium Rule

In the trapezium rule [50] we make the approximation

$$\int_a^b f(x)dx \approx h(\frac{1}{2}f_0 + f_1 + f_2 + \cdots + f_{n-2} + f_{n-1} + \frac{1}{2}f_n),$$

where

$$f_j = f(a + j\left(\frac{b-a}{n}\right)).$$

We could rewrite the trapezium rule as

$$\int_a^b f(x)dx \approx h\sum_{i=0}^{n}\omega_i f_i,$$

and obtain the weights ω_i by integrating the interpolating functions

$$\begin{array}{lll} p_1(t) = \frac{h-t}{h} & \text{for} & t \in [0,h], \\ p_k(t) = \frac{t-(k-1)h}{h} & \text{for} & t \in [(k-1)h, kh), \\ p_k(t) = 1 & \text{for} & t = k, \\ p_k(t) = \frac{(k+1)h-t}{h} & \text{for} & t \in [kh, (k+1)h), \\ p_n(t) = \frac{t-(n-1)}{h} & \text{for} & t \in [(n-1)h, nh]. \end{array} \tag{6.2}$$

6.3 Fractional Trapezium Rule

In [12] the quadrature weights are obtained by integrating the product of the interpolating functions (6.2) and the kernel of the fractional integral with the exception of the starting weight where the notion of a Hadamard finite part integral is additionally invoked to obtain the correct starting weight.

If in the integral

$$I_t = \int_0^t \frac{f(s)}{(t-s)^{1+\alpha}}ds, \tag{6.3}$$

we make the substitution $s = ut$ we obtain

$$I_t = t^{-\alpha}\int_0^1 \frac{f(ut)}{(1-u)^{1+\alpha}}du, \tag{6.4}$$

we may therefore, without loss of generality, derive the weights required for the unit interval divided in to j intervals of equal length h such that $h = 1/j$.

Interior Weights

To derive the interior weights for the fractional trapezium rule consider the integral I_k, $0 < k < j$, where

$$I_k = \frac{1}{h}\int_{(k-1)h}^{kh} \frac{(s-(k-1)h)}{(jh-s)^{1+\alpha}}ds + \frac{1}{h}\int_{kh}^{(k+1)h} \frac{((k+1)h-s)}{(jh-s)^{1+\alpha}}ds \qquad (6.5)$$

let $s = (j-u)h$ then we have $ds = -hdu$ and so

$$\begin{aligned} I_k &= -h^{-\alpha}\int_{j-(k-1)}^{j-k} \frac{(j-(k-1))-u}{u^{1+\alpha}}du - h^{-\alpha}\int_{j-k}^{j-(k+1)} \frac{u-(j-(k+1))}{u^{1+\alpha}}du \\ &= h^{-\alpha}\int_{j-k}^{j-(k-1)} \frac{(j-(k-1))-u}{u^{1+\alpha}}du - h^{-\alpha}\int_{j-(k+1)}^{j-k} \frac{(j-(k+1))-u}{u^{1+\alpha}}du \\ &= h^{-\alpha}[-\frac{(j-(k-1))}{\alpha u^{\alpha}} - \frac{u^{1-\alpha}}{1-\alpha}]_{j-k}^{j-(k-1)} - h^{-\alpha}[-\frac{(j-(k+1))}{\alpha u^{\alpha}} - \frac{-u^{1-\alpha}}{1-\alpha}]_{j-(k+1)}^{j-k} \end{aligned}$$

which gives

$$\begin{aligned} I_k &= h^{-\alpha}(-\frac{(j-(k-1))^{1-\alpha}}{\alpha} - \frac{(j-(k-1))^{1-\alpha}}{1-\alpha} \\ &+ \frac{(j-(k-1))}{\alpha(j-k)^{\alpha}} + \frac{(j-k))^{1-\alpha}}{1-\alpha} \\ &+ \frac{(j-(k+1))}{\alpha(j-k)^{\alpha}} + \frac{(j-k)^{1-\alpha}}{1-\alpha} \\ &- \frac{(j-(k+1))^{(1-\alpha)}}{\alpha} - \frac{(j-(k+1))^{1-\alpha}}{1-\alpha}) \end{aligned}$$

hence

$$\alpha(1-\alpha)h^{\alpha}I_k = 2(j-k))^{1-\alpha} - (j-(k-1))^{1-\alpha} - (j-(k+1))^{(1-\alpha)}. \quad (6.6)$$

End Weights

The singularity in the kernel occurs when $k = j$. However, for the test equation (6.1), we also know that the limit of the integral is finite, thus

$$\lim_{\epsilon \to 0} \int_0^{\epsilon} \frac{y(s)}{(\epsilon - s)^{\alpha}} ds = -\lambda y_0. \tag{6.7}$$

Consider the integral

$$I_{\epsilon} = \frac{1}{h} \int_{(j-1)h}^{jh-\epsilon} \frac{s - (j-1)h}{(jh - s)^{1+\alpha}} ds \tag{6.8}$$

let $s = (j - u)h$ then we have $ds = -hdu$ and so

$$\begin{aligned} h^{\alpha} I_{\epsilon} &= \int_{\epsilon}^{1} \frac{1-u}{u^{1+\alpha}} du \\ &= [-\frac{u^{-\alpha}}{\alpha} - \frac{u^{1-\alpha}}{1-\alpha}]_{\epsilon}^{1} \\ &= -\frac{1}{\alpha(1-\alpha)} + \frac{1}{\alpha \epsilon^{\alpha}} + \frac{\epsilon^{1-\alpha}}{1-\alpha}. \end{aligned}$$

We now appeal to the Hadamard finite part integral to conclude that the contribution of the convergent term to the weight is

$$\alpha(1-\alpha) h^{\alpha} I_j = -1, \tag{6.9}$$

and that the contribution of the divergent term to the required weight, when solving the test equation, will be y_0/α. This conclusion is justified by the assumption of a finite value for $y(0)$ in the test system.

For $k = 0$ we have

$$I_0 = \frac{1}{h} \int_0^{h} \frac{h-1}{(jh - s)^{1+\alpha}} ds \tag{6.10}$$

let $s = (j - u)h$ then we have $ds = -hdu$ and so

$$\begin{aligned} h^{\alpha} I_0 &= \int_{j-1}^{j} \frac{u - (j-1)}{u^{1+\alpha}} du \\ &= [\frac{u^{1-\alpha}}{1-\alpha} + \frac{(j-1)u^{-\alpha}}{\alpha}]_{j-1}^{j}, \end{aligned}$$

hence

$$\alpha(1-\alpha)h^{\alpha}I_0 = (\alpha-1)j^{-\alpha} - (j-1)^{1-\alpha} + j^{1-\alpha}. \tag{6.11}$$

The Convolution Weights

Writing the convolution weights as

$$\alpha(1-\alpha)h^{\alpha}\omega_{kj} = \begin{array}{ll} -1 & \text{for } k=0, \\ 2k^{1-\alpha} - (k-1)^{1-\alpha} - (k+1)^{1-\alpha} & \text{for } k=1,\ldots,j-1, \\ (\alpha-1)k^{-\alpha} - (k-1)^{1-\alpha} + k^{1-\alpha} & \text{for } k=j. \end{array} \tag{6.12}$$

To approximate the test system (6.1) $t = j/n$, $j, n \in \mathbb{N}$, we have

$$\frac{t^{-\alpha}}{\Gamma(-\alpha)}\left(\sum_{k=0}^{j}\omega_{kj}y_{j-k} + \frac{y_0}{\alpha}\right) = -\lambda y_j + f_j, \tag{6.13}$$

solving for y_j gives

$$(\omega_{0j} + t^{\alpha}\Gamma(-\alpha)\lambda)y_j = t^{\alpha}\Gamma(-\alpha)f_j - \sum_{k=1}^{j}\omega_{kj}y_{j-k} - \frac{y_0}{\alpha}. \tag{6.14}$$

Theorem 6.1 *([12] 2)*

Assuming that the functions involved are sufficiently smooth, the exists a constant η depending on α and y (and therefore on f and λ) such that the error in (6.14) is bounded by

$$|y(t_j) - y_j| \leq \eta j^{\alpha} n^{-2},\ j = 0, 1, \ldots, n. \tag{6.15}$$

Corollary 6.1 *([12] 3)*

If the functions involved are sufficiently smooth, we have the following global error estimate for (6.14):

$$\max_{j=0,1,\ldots,n} |y(t_j) - y_j| = \mathcal{O}(n^{\alpha-2}). \tag{6.16}$$

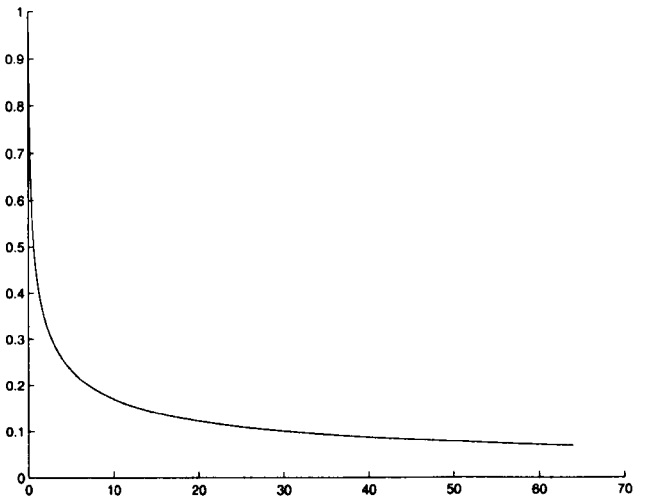

Figure 6.1: Approximate solution to (6.17).

Approximating the Test Equation

In Figure 6.1 we have plotted the approximate solution, calculated using the fractional trapezium rule, to the test equation with the specific representation

$$D^{0.5}y = -y,\ y(0) = 1, \tag{6.17}$$

with step length $h = 1/8$ over the interval $[0, 64]$. The number of Mflops used, by our MATLAB program, to produce this approximation was 0.40942.

In Figure 6.2 we have plotted the percentage error in the approximate solution, calculated using the fractional trapezium rule, to the test equation

$$D^{0.5}y = -y + t + \frac{t^{0.5}}{\Gamma(1.5)},\ y(0) = 0, \tag{6.18}$$

with step length $h = 1/8$ over the interval $[0, 64]$. The exact solution to (6.18) is $y(t) = t$.

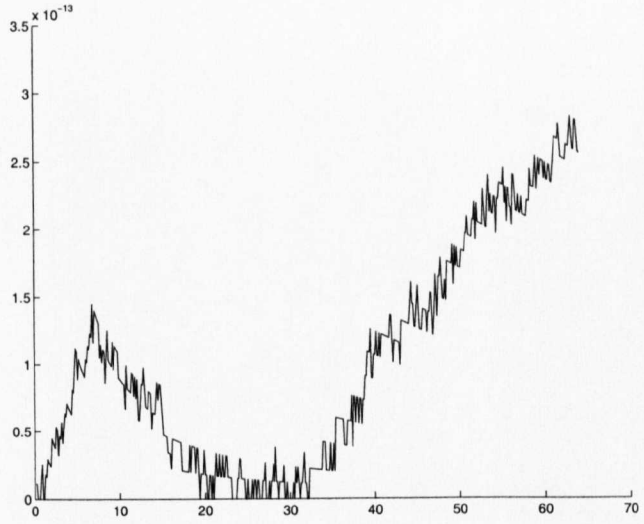

Figure 6.2: Percentage error in the approximate solution to (6.18).

Chapter 7

Fractional Linear Multistep Methods

7.1 Motivation

Lubich, in a sequence of papers [40, 41, 42, 43], develops a method for numerically approximating a convolution integral that combines the 19^{th} Century integration theory of Laplace and Cauchy with the theory of linear multistep methods we gave in Chapter 5. These methods are sometimes called convolution quadrature methods (CQM). The fractional integral is a type of convolution integral, and therefore CQM are candidate methods for numerically approximating the solutions of fractional differential equations.

In this chapter we review the application of CQM to the fractional calculus. CQM gives the appearance that the quadrature weights emanate directly from the convolution integral, as opposed to the more normal method of obtaining the quadrature weights by integrating a set of interpolating functions. This makes CQM seem the practical method of the greatest theoretical interest. However linear multistep methods play an essential part in obtaining the CQM quadrature weights, so a set of interpolating functions still ex-

ists in the background. In contradistinction to the fractional trapezium rule where we fractionally integrated a set of interpolating functions to obtain the weighs in CQM we use algebraic operations on characteristic polynomials of the backward differentiation formulae to generate our weights.

7.2 Convolution Integral and Inverse Laplace Transform Integrals

Consider the convolution integral

$$f * g(x) = \int_0^x f(x-t)g(t)dt, \; x \geq 0. \tag{7.1}$$

Let the kernel f be a function that is analytic and exponentially bounded, on a sector containing the positive real half line, as $t \to \infty$ and $f(t) = O(t^{\mu-1})$, where $\mu > 0$, as $t \to 0$. Equivalently, we could require that the Laplace transform of f, written $F(s)$, exists and is analytic in a sector $|\arg(s-c)| < \pi - \phi$ with $\phi < \frac{\pi}{2}$, $c \in \mathbb{R}$ and

$$|F(s)| \leq M.|s|^{-\mu}, \;\; M < \infty, \; \mu > 0. \tag{7.2}$$

Then, by using the Laplace inversion integral, we can write f as

$$f(t) = \frac{1}{2\pi i} \int_\Gamma F(\lambda) e^{\lambda t} d\lambda, \; t > 0, \tag{7.3}$$

where Γ is a contour running from $\infty.e^{-i(\pi-\phi)}$ to $\infty.e^{i(\pi-\phi)}$ within a sector where $F(s)$ is analytic and enclosing all the poles of $F(s)$, see [41].

Substituting this representation for f into (7.1) gives

$$\int_0^x f(x-t)g(t)dt = \frac{1}{2\pi i} \int_\Gamma F(\lambda) \int_0^x e^{\lambda t} g(x-t)dt. \tag{7.4}$$

Now the differential equation

$$y' = \lambda y + g,\ y(0) = y_0,$$

has as its solution

$$y = y_0 e^{\lambda t} + \int_0^x e^{\lambda t} g(x-t)dt. \tag{7.6}$$

We set $y_0 = 0$ and then use a LMM to solve for y in (7.6) and substitute this in (7.4).

Thus we have

$$\sum_{j=0}^{k} \alpha_j y_{n+j-k} = h \sum_{j=0}^{k} \beta^j (\lambda y_{n+j-k} + g((n+j-k)h)),\ n \geq 0, \tag{7.7}$$

with starting values $y_{-k} = \ldots y_{-1} = 0$, and with $g \in C[0,\infty)$ extended by 0 to the negative real axis, i.e. $g(t) \equiv$ for $t < 0$. Multiplying (7.7) by ζ^n and summing over n from 0 to ∞ we obtain

$$(\alpha_0 \zeta^k + \cdots + \alpha_k).\mathbf{y}(\zeta) = (\beta_0 \zeta^k + \cdots + \beta_k).(h\lambda.\mathbf{y}(\zeta) + h.\mathbf{g}(\zeta)), \tag{7.8}$$

where $\mathbf{y}(\zeta) = \sum_{n=0}^{\infty} y_n \zeta^n$, and $\mathbf{g}(\zeta) = \sum_{n=0}^{\infty} g(nh)\zeta^n$ are formal generating power series.

Let $\delta(\zeta) = (\alpha_0 \zeta^k + \cdots + \alpha_{k-1}\zeta + \alpha_k)/(\beta \zeta^k + \cdots + \beta_{k-1}\zeta + \beta_k)$ then y_n is the nth coefficient of the formal power series $(\delta(\zeta)/h - \lambda)^{-1}\mathbf{g}(\zeta)$.

So (7.4) can be approximated at $x = nh$ by the nth coefficient of

$$\frac{1}{2\pi i} \int_\Gamma F(\lambda)(\frac{\delta(\zeta)}{h} - \lambda)^{-1}\mathbf{g}(\zeta)d\lambda = F(\frac{\delta(\zeta)}{h})\mathbf{g}(\zeta) \tag{7.9}$$

with equality holding by virtue of Cauchy's integral formula [3]. Writing

$$F(\frac{\delta(\zeta)}{h}) = \sum_{j=0}^{\infty} \omega_j(h)\zeta^j \tag{7.10}$$

then the coefficients of the right hand side of (7.9) are the Cauchy product of the two sequences $\{\omega_j(h)\}$ and $\{g(jh)\}$.

We assume that the LMM is $A(\alpha)$ stable with $\alpha > \phi$ with ϕ as in (7.2), stable in a neighbourhood of infinity, strongly zero-stable and consistent of order p. Equivalently as conditions on $\delta(\zeta)$ we require that $\delta(\zeta)$ is analytic and without zeros in a neighbourhood of the closed unit disc $|\zeta| \geq 0$, with the exception of a zero at $\zeta = 1$,

$$|\arg \delta(\zeta)| \leq \pi - \alpha, \quad |\zeta| < 1, \text{ for some } \alpha > \phi, \tag{7.11}$$

$$\frac{1}{h}\delta(e^{-h}) = 1 + O(h^p), \text{ for some } p \geq 1. \tag{7.12}$$

Now the fractional integral is given by

$$(I^{\alpha}_{0+}\phi) = \frac{1}{\Gamma(\alpha)} \int_0^x \frac{\phi(t)}{(x-t)^{1-\alpha}} dt, \; x > a, \tag{7.13}$$

which has as kernel, up to an arbitrary multiplicative factor,

$$f(x) = \frac{x^{\alpha-1}}{\Gamma(\alpha)}, \quad 0 < \alpha < 1. \tag{7.14}$$

The Laplace transform of f is $s^{-\alpha}$ consequently the weights we require $\{\omega_n(h)\}$ are the coefficients of ζ^n in the expansion of $h^{\alpha}\delta(\zeta)^{-\alpha}$.

Starting Weights

Using the notation in [42] for the fractional quadrature at $t = nh$ we have

$$\Omega^{\alpha}_h f(x) = h^{\alpha} \sum_{j=0}^{n} \omega_{n-j} f(jh). \tag{7.15}$$

We write the convolution quadrature error operator as E^{α}_h so for a suitably integrable function y we have

$$E^{\alpha}_h y = \Omega^{\alpha}_h y - I^{\alpha} y. \tag{7.16}$$

To obtain the order p convergence of the underlying multistep method we choose a set of polynomials $q_i(t)$, $i = 1, \ldots, p$, such as t^β, $t^{\beta+1}, \ldots, t^{\beta+p-1}$, and calculate a set of starting weights w_{nj} so that the $E_h^\alpha q_i(nh) = 0$, $i = 1, \ldots, p$.

The particular set of polynomials that we should require the quadrature method to be exact for depends upon which fractional differential operator we wish to approximate. $E_h^\alpha y = \sum_{j=0}^{n} w_{nj} y(jh)$ is called the starting quadrature.

Theorem 7.1 *([40] 139) Under the assumptions (7.2) for a LMM satisfying the conditions given above, the method in (7.16), for $y \in C^p[0,t]$, has $\mathcal{O}(h^p)$ convergence, i.e.*

$$|y(t_n) - y_n| \leq C t_n^{\mu-1} h^p, \tag{7.17}$$

with $C > 0$ independent of $h > 0$ and $t_n \geq h$.

To approximate the Caputo fractional derivative we have

$$\sum_{j=0}^{\infty} \omega_j(h) \zeta^j = \frac{\delta(\zeta)}{h} (\frac{\delta(\zeta)}{h})^{\alpha-1} = (\frac{\delta(\zeta)}{h})^\alpha.$$

To obtain the $\mathcal{O}(h^p)$ convergence of the underlying LMM we calculate the starting weights so that the quadrature is exact for the set of power functions $1,\ t,\ \ldots\ t^{p-1}$.

For other fractional differential operators the starting quadrature weights w_{nj} can be determined (by theorem 2.4 in [42] see also [7]) as follows: fix $\beta \neq -1, -2, -3, \ldots$, let m be an integer such that $\mathrm{Re}(m + \beta - 1) < p \leq \mathrm{Re}(m + \beta)$ and put

$$h^\alpha \sum_{j=1}^{m} w_{nj} (jh)^{q+\beta-1} + (E_h^\alpha t^{q+\beta-1})(1) = 0. \tag{7.18}$$

Then multiplying by $n^{q+\alpha+\beta-1}$ and rearranging with respect to (7.16) gives

$$\sum_{j=1}^{m} w_{nj} j^{q+\beta-1} = \frac{\Gamma(q+\beta)}{\Gamma(q+\beta+\alpha)} n^{q+\alpha+\beta-1} - \sum_{j=1}^{n} \omega_{n-j} j^{q+\beta-1}, \tag{7.19}$$

where $q = 0, \ldots, m-1$. This ensures that the starting approximation is exact for polynomials up to degree $p-1$.

7.2.1 Determination of $\{\omega_n\}$

The reciprocal, logarithm and exponential of a polynomial or formal power series can be calculated by using fast Fourier transforms (FFT) [29]. This is because the convolution properties of FFT's provide a relatively economical algorithm for performing calculations of this kind. A method for calculating the $\{\omega_n\}$ required for fractional linear multistep methods is given in [5, 6, 7]. To obtain the coefficients of $\delta(\zeta)^{-\alpha}$ FFT transforms are used to calculate the coefficients in the expansion of $\exp(\alpha \log(1/\delta(\zeta))$.

The mathematics in the presentation in [5, 6, 7] is entangled with FORTAN programming. MATLAB frees us from the obfuscation engendered by programming considerations we can therefore present the method in [5, 6, 7] concentrating purely on the mathematics.

Calculation of the Reciprocal of $p(x)$

Let $p(x) = \alpha_0 + \cdots + \alpha_n x^n$ where $\alpha_0 \neq 0$ and let $1/p(x) = r(x)$. To calculate $r(x)$ we use Newton-Raphson iteration. Let

$$f = \frac{1}{r} - p \tag{7.20}$$

we wish to find the zero of f. To do this we use the Newton-Raphson iteration $r_{n+1} = r_n - f_n/f'_n$. So with

$$\frac{f}{f'} = \frac{\frac{1}{r} - p}{\frac{-1}{r^2}},$$

we therefore have the recurrence

$$r_{n+1} = r_n(2 - r_n p). \tag{7.21}$$

The use of FFTs doubles the number of correct coefficients with each iteration.

Calculation of the Log of $p(x)$

Let $q(x) = (\alpha_1 x + \cdots + \alpha_n x^n)/\alpha_0$ where $\alpha_1 \neq 0$ (such a polynomial is called a non-unit in the terminology of [5, 6, 7]) then

$$\log(p(x)) = \log(\alpha_0 + \alpha_0 q(x)) = \log(\alpha_0) + \log(1 + q(x)). \tag{7.22}$$

We calculate the log of $p(x)$ by using the integral equality

$$\log(p(x)) = \log(\alpha_0) + \int_0^x \frac{q'(x)}{1 + q(x)} dx = \log(\alpha_0) + \int_0^x \frac{p'(x)}{p(x)} dx. \tag{7.23}$$

Therefore by calculating the reciprocal of p then multiplying it by p' and integrating the resulting product we can obtain $\log(p(x))$ in three steps.

Calculation of the Exponential of $p(x)$

Using the notation above

$$\exp(p(x)) = \exp(\alpha_0 + q(x)) \tag{7.24}$$

where

$$\exp(q(x)) = 1 + q(x) + \frac{q(x)^2}{2!} + \frac{q(x)^3}{3!} + \ldots. \tag{7.25}$$

If $\epsilon(x) = \exp(q(x)) - 1$ then $\epsilon(x)$ is a non-unit and since $\log(1+\epsilon(x)) = q(x)$ we can solve this equation, by mean of the Newton-Raphson iteration, for $\epsilon(x)$ and then $\exp(p(x)) = \exp(\alpha_0)(1 - \epsilon(x))$.

7.2.2 The FLMM Approximation of the Caputo Fractional Derivative

We have used fractionalised BDF1 and BDF2 to approximate the 0.9 order derivative of the power function t and plotted the function values in Figure 7.1 and the percentage error in Figure 7.2. In Figure 7.1 the lower curve is the approximation calculated with BDF1 and the upper curve is the calculation calculated with BDF2 the circles being the exact values of the derivative. In Figure 7.2 the upper curve is the error for BDF1 and the lower curve is for BDF2. In fact the lower curve is identically zero since the starting weights are constructed to make the approximation exact for linear functions.

Approximating the Test Equation

When using fractionalised BDF1 the starting weight required at each step is just the additive inverse of the sum of the fractional linear multistep weights used up to and including that step. This makes the method easy to implement as can be seen from the program listing.

```
% LDL01full
% Lubich's method on Caputo's derivative
clear;
p=[1,-1];    % defines Euler's method or BDF1
% set the order of derivative, time step, integration interval etc.
```

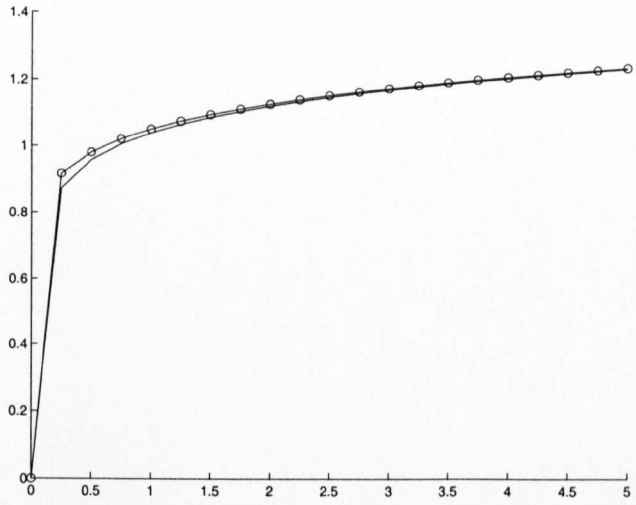

Figure 7.1: $D^{0.9}t$ using BDF1 and BDF2.

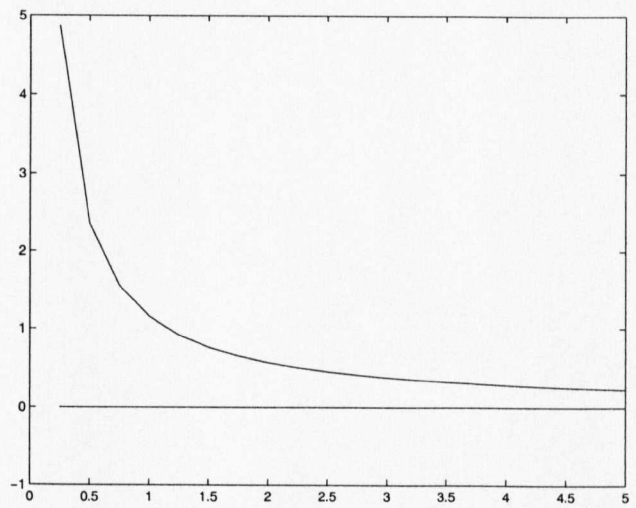

Figure 7.2: The percentage error in the $D^{0.9}t$ using BDF1 and BDF2.

```
LDL_set_parameters
LDL_do_weights % calculates the weights
%                        do the integration
start_weight1=-weights(1);
for i=2:1:steps
   start_weight1(i)=start_weight1(i-1)-weights(i);
end
for i=2:1:steps
   y_sum=start_weight1(i)*y(1)+weights(2:i)*y(i-1:-1:1)';
   y(i)=front*(f(i)-y_sum);
end
LDL_plot   % plot the result
```

In Figure 7.3 we have plotted the approximate solution, calculated using fractionalised BDF1, to the test equation

$$D^{0.5} = -y,\ y(0) = 1, \tag{7.26}$$

with step length $h = 1/8$ over the interval $[0, 64]$.

The number of Mflops used to produce this approximation, by our program in MATLAB, was 34.42021. The difference in computational effort between FLMM and fractional trapezium rule is due to the computation of the weights. The weights for fractional BDF1 were calculated using the procedure outlined above which is mathematical overkill in this particular case but does suggest a difficulty when using a LMM with more terms.

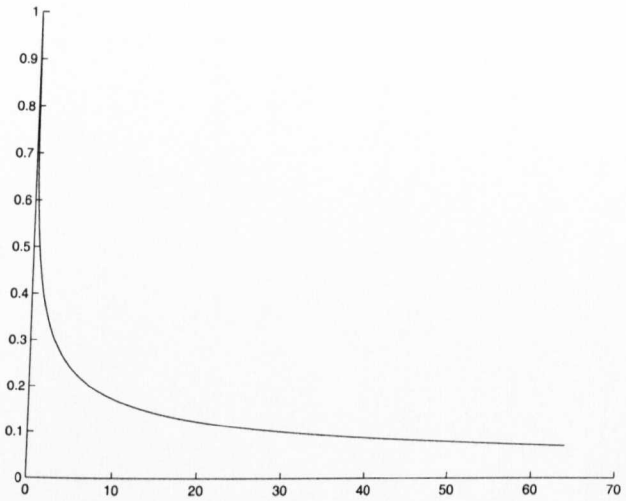

Figure 7.3: Approximate solution to (7.26).

In Figure 7.4 we have plotted the percentage error in the approximate solution, calculated using the fractionalised BDF1, to the test equation

$$D^{0.5} = -y + t + \frac{t^{0.5}}{\Gamma(1.5)}, \; y(0) = 0, \tag{7.27}$$

with step length $h = 1/8$ over the interval $[0, 64]$.

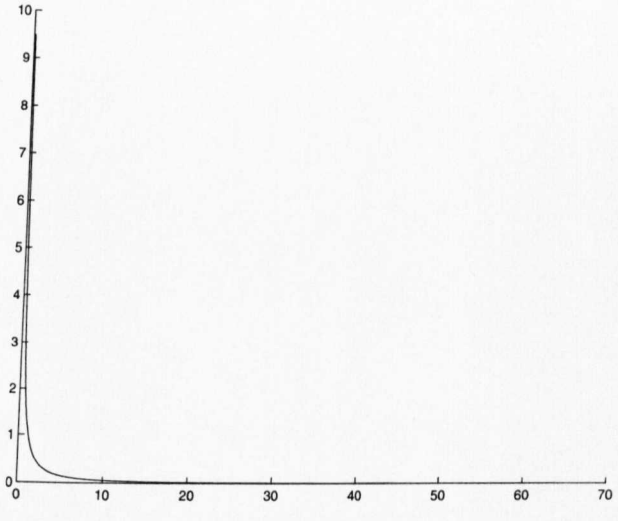

Figure 7.4: Percentage error in the approximate solution to (7.27).

Chapter 8

Finite Memory and Sparse Quadrature

8.1 Motivation

In this chapter we describe Podlubny's Finite Memory Principle [48, 49] and Sloan and Thomee's Sparse Quadrature [56]. These are both attempts to reduce the computational effort required to calculate an approximate solution to a fractional differential equation.

The methods for approximating the solution of fractional differential equations, given in the previous chapters were concerned with accurately approximating the singularity at the origin in the kernel function. This is essentially a mathematical problem. The second computational problem posed in the approximation of the solution of a fractional differential equation, the $\mathcal{O}(n^2)$ computational cost of approximating the convolution integral of two arbitrary functions, remains.

In the classical calculus time is valued linearly and therefore when approximating the solution of an ordinary differential equation the relation of each past *moment of time* to the present remains the same. Due to the non-

linear valuation of time in the fractional calculus at each iteration we have to re-evaluate the relation of each past *moment in time* to the present to determine the behaviour of the solution.

When approximating the solution of an ordinary differential equation at time $t = ih$ the calculation takes $\mathcal{O}(1)$ flops. For a fractional differential equation the equivalent calculation takes $\mathcal{O}(i)$ flops when using the fractional trapezium rule or a FLMM. Thus the computational effort required to approximate the solution to an ordinary differential equation over an interval $[0, ih]$ grows linearly with i, whilst the computational effort required to approximated the solution to a fractional differential equation grows quadratically with i. Here we examine two attempts to reduce the computational cost of approximating the solutions to fractional differential equations.

8.2 Podlubny's Finite Memory Principle

Podlubny [48, 49] suggests that the computational effort required to approximate the solution to a fractional differential equation may be reduced to linear growth by limiting the interval of integration, in the convolution integral, to a window of length $T > 0$. Therefore instead of integrating over the whole of $[0, t]$ we integrate over a moving window of fixed length, i.e. over the interval $[t - T, t]$.

This introduces into the approximation of the fractional derivative an error E_t given by

$$E_t = |\frac{1}{\Gamma(1-\alpha)} \int_0^{t-T} \frac{y'(s)}{(t-s)^\alpha} ds|. \tag{8.1}$$

Let $M = \max_{t\geq 0} |y(t)|$ then integrating by parts gives

$$E_t \leq \frac{1}{\Gamma(1-\alpha)}(|[\frac{y(s)}{(t-s)^\alpha}]_0^{t-T}| + \alpha|\int_0^{t-T} \frac{y(s)}{(t-s)^{1+\alpha}}ds| \tag{8.2}$$

and therefore

$$E_t \leq \frac{2M}{T^\alpha \Gamma(1-\alpha)}. \tag{8.3}$$

As we can see this error bound is independent of the step length h therefore the error due to the finite memory principle can not be controlled by altering the step length.

The usefulness of the finite memory principle depends critically on the behaviour of the solution y. If y is unbounded then the error may become unbounded. This is clearer if we rewrite the upper limit in the truncation formula (8.1) as $t(1 - T/t)$ then as $t \to \infty$ the truncation becomes almost the whole of the fractional derivative since $T/t \to 0$.

As an example consider the integral occurring in the *beta function.* For $p > 0$ and $q > 0$ the *beta function* is given by

$$\int_0^1 x^{p-1}(1-x)^{q-1} = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \tag{8.4}$$

Substituting $st = x$ and writing $1 - q = \alpha$ and $p = \beta$ gives

$$\int_0^t \frac{s^{\beta-1}}{(t-s)^\alpha}ds = \frac{\Gamma(\beta)\Gamma(1-\alpha)}{\Gamma(\beta-\alpha+1)}t^{\beta-\alpha}. \tag{8.5}$$

Clearly unless $y(t)$ is a rapidly decreasing function, i.e. $\mathcal{O}(t^{\beta-1})$ where $0 < \beta < 1$, we would expect an increasing error to occur.

If y is not rapidly decreasing then to reduce the error we will need to increase the interval of integration to such an extent that no practical reduction in the amount of computational effort required to calculate the approximate solution is obtained.

8.2.1 Computational Assays

We have applied the finite memory principle to the fractional trapezium rule to assess its affect on the approximate solution. The effect of applying the finite memory principle, in comparison with the solution calculated over the full interval to the system,

$$D^{\alpha}y(t) = -y(t) + f(t),\ 0 < \alpha < 1, \tag{8.6}$$

is shown in figures 8.1 and 8.2. Here we have used a step length of 0.125. In both cases the solid curve is the solution calculated by integrating over the full interval $[0, t]$ and the dotted curve is the approximate solution calculated by integrating over a finite window of fixed length $[t-8, t]$, for $t > 8$.

In figure 8.1 for $f(t) \equiv 0$ and $y(0) = 1$ the computation required for the finite memory calculation was $\mathcal{O}(0.073)$ Mflops compared to $\mathcal{O}(0.285)$ Mflops required for the full memory. This confirms that the reduction in computational effort takes place however as we can see a divergence between the solutions rapidly occurs.

In Figure 8.2 we have used $f = t + t^{0.5}/\Gamma(1.5)$ and $y(0) = 0$, the exact solution in this case is $y = t$ i.e. a straight line. We see again the immediate onset of an increasing error.

8.3 Sparse Quadrature

In [56] Sloan and Thomee suggest reducing the computational effort required to approximate the solution to an integro-differential equation, of which fractional differential equations form a subclass, by dividing the interval of integration, into two intervals, and using approximation methods of different

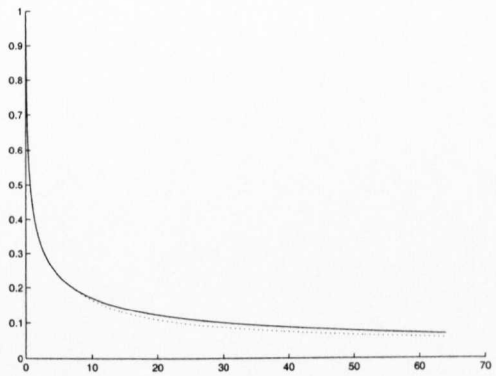

Figure 8.1: Comparison for $f(t) \equiv 0$ and $y(0) = 1$.

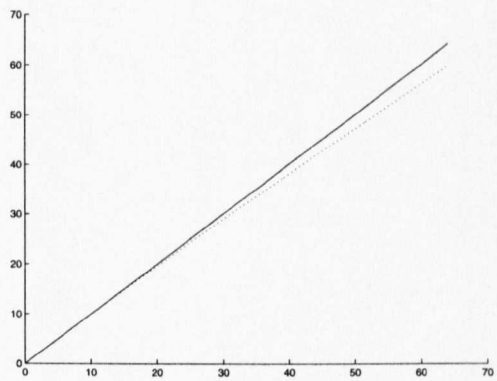

Figure 8.2: Comparison for $f(t) = t + t^{0.5}/\Gamma(1.5)$ and $y(0) = 0$.

orders in each of the intervals to obtain an approximate solution. We attempt to describe this method in terms appropriate to our problem.

Let the step length h be given. Choose some $m > 0$, let $T = mh$ and let l be the largest integer strictly less than n/m. Then at time $t = nh > T$ let $T_1 = [lmh, nh]$ and $T_2 = [0, lmh]$ we can then divide the integral of integration thus

$$[0, t] = T_2 \cup T_1. \tag{8.7}$$

In *sparse quadrature* we calculate an approximate solution by using step length h in T_1 and mh in T_2.

In [56] it is suggested that we chose $m = h^{-1/2}$ and for T_1 we use an approximation method of order p and in T_2 we use an approximation method of order $2p$. Thus the order of the approximation calculated using the *sparse quadrature method will remain p.*

Clearly whilst this method could reduce the computational effort required it will still exhibit quadratic growth in computational effort at the rate of the second method, i.e. we replace $\mathcal{O}(i^2)$ with $\mathcal{O}((i/m)^2)$ which is the same as $\mathcal{O}(i^2)$ since m is a constant.

Chapter 9

Acceleration by Extrapolation

9.1 Motivation

In Richardson extrapolation two or more approximations are exploited to produce a more accurate approximation. The use of Richardson extrapolation with the trapezium rule is know as Romberg integration [50]. Application of extrapolation to the trapezium rule enables a more accurate approximate solution to be obtained from the previously calculated values of the approximate solution for little additional computational effort. In this chapter we present Diethelm and Walz's [17] application of Richardson extrapolation to the fractional trapezium rule to produce a fractionalised version of Romberg integration. Here also the application of Richardson extrapolation enables a more accurate approximation of the true solution to be obtained with little further computational effort.

9.2 Richardson Extrapolation

We use the definition of Richardson extrapolation given in [50]. Suppose we have an approximating functional $\phi = \phi(f(t))$ of the function f. We can

generate a set of approximations ϕ_i

$$\phi_i(f(t)) = \phi(f(t) : h_i),\ i = 1, 2, \ldots,$$

where the step length h_i is such that $i > j \longleftrightarrow h_i < h_j$ and $h_i \to 0$ as $i \to \infty$. If the error has the asymptotic form

$$E = \sum_{j=1}^{\infty} a_j h_i^{\gamma_j},\ 0 < \gamma_1 < \gamma_2 < \ldots, \tag{9.1}$$

where the constants a_j are independent of the step length h_i, we can write the true value of ϕ as

$$\phi = \phi_i + \sum_{j=1}^{\infty} a_j h_i^{\gamma_j},\ i = 1, 2, \ldots. \tag{9.2}$$

Let $i = 1$ in (9.2) and multiply by $h_2^{\gamma_1}$ and similarly let $i = 2$ and multiply by $h_1^{\gamma_1}$ and then subtract the resulting equations and solve for ϕ we obtain

$$\phi = \frac{1}{h_1^{\gamma_1} - h_2^{\gamma_1}}(h_1^{\gamma_1}\phi_2 - h_2^{\gamma_1}\phi_1) + \sum_{j=2}^{\infty} a_j \frac{h_1^{\gamma_1} h_2^{\gamma_j} - h_2^{\gamma_1} h_1^{\gamma_j}}{h_1^{\gamma_1} - h_2^{\gamma_1}}. \tag{9.3}$$

Now let $h_2 = \rho h_1,\ 0 < \rho < 1$ in (9.3), by cancellation of powers of h_1, we obtain

$$\phi = \frac{\phi_2 - \rho^{\gamma_1}\phi_1}{1 - \rho^{\gamma_1}} + \sum_{j=2}^{\infty} \frac{\rho^{\gamma_j} - \rho^{\gamma_1}}{1 - \rho^{\gamma_1}} a_j h_1^{\gamma_j}. \tag{9.4}$$

If we let

$$\phi_{12} = \frac{\phi_2 - \rho^{\gamma_1}\phi_1}{1 - \rho^{\gamma_1}} \quad \text{and} \quad b_j = \frac{\rho^{\gamma_j} - \rho^{\gamma_1}}{1 - \rho^{\gamma_1}} a_j$$

we can rewrite (9.4) as

$$\phi = \phi_{12} + \sum_{j=2}^{\infty} b_j h_i^{\gamma_j}. \tag{9.5}$$

Similarly if we were to repeat the operation given above with $h_3 = \rho h_2$ we could compute a ϕ_{23} and then eliminate the $j = 2$ term in (9.5) to obtain

ϕ_{123}. Continuing this process recursively we can generate an array of the form

$$\begin{array}{ccccc} \phi_1 & & & & \\ & \phi_{12} & & & \\ \phi_2 & & \phi_{123} & & \\ & \phi_{23} & & \phi_{1234} & . \\ \phi_3 & & \phi_{234} & & \\ & \phi_{34} & & & \\ \phi_4 & & & & \end{array} \tag{9.6}$$

9.2.1 Romberg Integration

The trapezium rule is a particular case where the functional ϕ possesses a suitable asymptotic error expansion. By the Euler-Maclaurin sum formula [50], for an integrable function f, we have

$$\begin{aligned} \sum_{j=0}^{m} f(a+jh) &= \frac{1}{h}\int_a^b f(y)dy(y) + \frac{1}{2}(f(a)+f(b)) \\ &\quad + \sum_{k=1}^{m} \frac{B_{2k}}{(2k)!}(f^{(2k-1)}(b) - f^{(2k-1)(a)}) + E_m \end{aligned}$$

where the B_{2k} are Bernoulli numbers, the error has the form

$$E_m = \frac{nh^{2m+2}B_{2m+2}}{(2m+2)!} f^{(2m+2)}(\zeta), \tag{9.7}$$

and $h = (b-a)/m$. From this we can write the trapezium rule as

$$\int_a^b f(x)dx = h(\frac{1}{2}f_0 + f_1 + f_2 + \cdots + f_{m-1} + \frac{1}{2}f_m) + \sum_{j=1}^{\infty} a_j h^{2j}. \tag{9.8}$$

Therefore the trapezium rule has an asymptotic expansion of a form which is suitable for the use of Richardson extrapolation.

9.3 Extrapolation and Fractional Differential Equations

In [17] it is shown that the fractional trapezium rule possesses an asymptotic error expansion such that we can apply Richardson extrapolation to obtain a sequence of improved approximate solutions.

Suppose we have used the fractional trapezium rule on the fractional differential equation

$$D^{\alpha}y = \lambda y + f,\ y(0) = y_0\ \lambda < 0,\ 0 < \alpha < 1, \tag{9.9}$$

to generate the approximate solution $\{x_i : i = 0, \ldots, n\}$.

Theorem 9.1 *([17] 231)*

There exist coefficients $c_\mu = c_\mu(\alpha)$ *and* $c^*_\mu = c^*_\mu(\alpha)$ *such that the sequence* $\{y_n\}$ *possess an asymptotic expansion of the form*

$$y_n = y(t_n) + \sum_{\mu=2}^{M_1} c_\mu n^{\alpha-\mu} + \sum_{\mu=1}^{M_2} c^*_\mu n^{-2\mu} + o(n^{-M_3}) \quad for\ n \to \infty, \tag{9.10}$$

where M_1 *and* M_2 *depend on the smoothness of* x *and* f, *and* $M_3 = \min\{\alpha - M_1, 2M_2\}$.

9.3.1 Implementation

To develop an algorithm it is convenient to rewrite (9.10) as

$$y_n = y(t_n) + \sum_{\mu=1}^{M} \gamma_\mu n^{-\lambda_m u} + o(n^{-\lambda_M}) \tag{9.11}$$

for $n \to \infty$ where $j = 1, 2, \ldots,$ we have

$$\lambda_{3j} = 2j + 1 - \alpha,\ \ \lambda_{3j-1} = 2j,\ \ \lambda_{3j-2} = 2j - \alpha. \tag{9.12}$$

Choose integers $n_0 > 0$, $b > 1$ (usually we let $b = 2$), and $0 < K \leq M$ such that for $i = 0, 1, 2, \ldots$, we have $t_n = n_i h_i$ where $h_i = 1/(n_0 b^i)$. Next compute the sequence of approximations $x_i^{(0)} = y_{n_i}$ at t_n.

We are now in a position to apply Richardson extrapolation as given above using the formula

$$x_i^{(k)} = x_{i+1}^{(k-1)} + \frac{x_{i+1}^{(k-1)} - x_i^{(k-1)}}{b^{\lambda_k} - 1}, \tag{9.13}$$

for $k = 1, 2, \ldots, K$ and $i = 0, 1, \ldots$.

Theorem 9.2 *([17] 231)*

For arbitrary K the use of (9.13) with the fractional trapezium rule is stable.

We now give an example which illustrates the convergence of the extrapolation scheme. We will use an initial step length of 0.5 and we let $b = 2$ and $K = 5$.

Firstly, in table 9.1, we give the fractional Romberg tableau for the system

$$D^{0.5}y(t) = -y(t) + t^2 + \frac{2t^{1.5}}{\Gamma(2.5)}, \; y(0) = 0, \tag{9.14}$$

at $t = 16$.

The total number of flops used to calculate the approximate solutions and the consequent Richardson extrapolation was 489192. To calculate an approximation to $x(16)$ using a step length of half the shortest length used in the Romberg scheme requires 1.1941 Mflops and has an error of 7.85 (-4). Therefore the extrapolation scheme compares very favourably with 'brute' force computation.

Step	flops	error $o(n^{\alpha-2})$	error $o(n^{-2})$	error $o(n^{\alpha-3})$	error $o(n^{\alpha-4})$	error $o(n^{-4})$
0.5	5556	$-1.37(-1)$				
			$-9.37(-4)$			
0.25	13140	$-4.91(-2)$		$2.56(-6)$		
			$-2.32(-4)$		$-5.70(-8)$	
0.125	34452	$-1.75(-2)$		$4.07(-7)$		$-1.20(-9)$
			$-5.77(-5)$		$-6.13(-9)$	
0.0625	101652	$-6.23(-3)$		$6.69(-8)$		
			$-1.43(-5)$			
0.03125	334356	$-2.21(-3)$				

Table 9.1: Fractional Romberg tableau for system (9.14).

We conclude that for good initial starting conditions the extrapolation scheme in [17] can be used to obtain a considerable improvement in the convergence of the approximate solution to the true solution at the interpolation points of the first approximation.

We have truncated the reporting of the error in the table to three significant digits and we have rewritten the usual scientific notation for numbers $a \times 10^{-b}$ as $a(-b)$ so that we can fit all the data on one page. Further examples are given in [17].

Chapter 10

Acceleration by Recursive Summation

10.1 Motivation

In this chapter we give a new interpretation of the fractional integral. We use the scaling property of the fractional integral to express it as the sum of integrals over the same interval. This enables us to implement the numerical methods we have given for approximating the solution of a fractional differential equation in such away that we require only $\mathcal{O}(n \log n)$ arithmetic operations as n increases.

We first give an algorithm with the *least cost* implementation of our ideas, which we will call the *Nested Mesh* version, which, although it involves a slight truncation of the integral, should have sufficient accuracy for enough practical applications, such as embedded real time controllers, to justify its presentation. Then we give an algorithm, which involves no truncation of the integral, which we will call the *Fractal Sum* version. In both cases we demonstrate the practicality of our method by showing how it can be implemented using the fractional trapezium rule.

In the fractional trapezium rule a fixed step length is used throughout the whole of the interval of integration. We could increase the size of the step length by using a polynomially graded mesh however it has been shown that this approach can lead to stability problems [13].

Our approach will be to use a given step length for the most recent part of the integral and then to use progressively longer step lengths in the earlier parts of the integral in a systematic way.

10.2 The Fractional Derivative as an Infinite Sum of Integrals

The fractional differential operator possesses a scaling property which gives it a relation to the theory of fractals. This is also known as the homogeneity property of the fractional integral by some authors [40], see also [30]. We will now use this scaling property to express the fractional derivative as a sum of integrals over the same interval. This interval then provides us with an invariant scale.

Let $\alpha \in (0,1)$, for a differentiable function y the Caputo derivative at $t = 2^n$ is:

$$D^{\alpha}y(t)|_{t=2^n} = I^{1-\alpha}Dy(t)|_{t=2_n} = \frac{1}{\Gamma(1-\alpha)}\int_0^{2^n} \frac{y'(2^n - s)}{s^{\alpha}}ds. \qquad (10.1)$$

We rewrite (10.1) as a sum of integrals over the intervals $[2^{i-1}, 2^i]$, $i \in \mathbb{Z}$, $i \leq n$, which gives

$$D^{\alpha}y(t)|_{t=2^n} = \frac{1}{\Gamma(1-\alpha)}\sum_{i=-\infty}^{n}\int_{2^{i-1}}^{2^i} \frac{y'(2^n - s)}{s^{\alpha}}ds. \qquad (10.2)$$

We now rescale these integrals to be integrals over the same interval $[0.5, 1]$

by making the substitutions $s \to 2^i u$, $ds \to 2^i du$, $y'(s) \to y'(2^i u)/2^i$, to give

$$D^\alpha y(t)|_{t=2^n} = \frac{1}{\Gamma(1-\alpha)} \sum_{i=-\infty}^{n} \frac{1}{2^{i\alpha}} \int_{\frac{1}{2}}^{1} \frac{y'(2^n - 2^i u)}{u^\alpha} du = \sum_{i=-\infty}^{n} D_i. \qquad (10.3)$$

This gives an expression where the singularity at the origin is replaced by a convergent infinite sum.

For negative i of large magnitude we will require values of $y'(t)$ for values of t far less than any practical step length, so we choose some initial 2^{i_0} such that

$$D^\alpha y(t)|_{t=2^n} = \frac{1}{\Gamma(1-\alpha)} \frac{1}{2^{i_0\alpha}} \int_0^{\frac{1}{2}} \frac{y'(2^n - 2^{i_0} u)}{u^\alpha} du + \sum_{j=i_0+1}^{n} D_j. \qquad (10.4)$$

Let

$$D_{i_0} = \frac{1}{\Gamma(1-\alpha)} \frac{1}{2^{i_0\alpha}} \int_0^{\frac{1}{2}} \frac{y'(2^n - 2^{i_0} u)}{u^\alpha} du, \qquad (10.5)$$

then

$$D^\alpha y(t)|_{t=2^n} = D_{i_0} + \sum_{j=i_0+1}^{n} D_j. \qquad (10.6)$$

We now conclude that if y and y' are sufficiently well behaved then we can solve each of the integrals in (10.6) with the same step length, the convergence following from the results for the underlying method used for obtaining the approximate solution.

10.2.1 Writing the convolution the other way round

We can rewrite the previous as follows, let $\alpha \in (0,1)$ then for a differentiable function y we can write the Caputo derivative on $[0, 2^n]$ as:

$$D^\alpha y(t)|_{t=2^n} = I^{1-\alpha} D y(t)|_{t=2_n} = \frac{1}{\Gamma(1-\alpha)} \int_0^{2^n} \frac{y'(s)}{(2^n - s)^\alpha} ds. \qquad (10.7)$$

Let

$$D_0 = \frac{1}{\Gamma(1-\alpha)} \int_{2^{n-1}}^{2^n} \frac{y'(s)}{(2^n - s)^\alpha} ds, \tag{10.8}$$

then we can write (10.7) as

$$D^\alpha y(t)|_{t=2^n} = D_0 + \sum_{i=1}^{n-1} \int_{2^{n-2^{i+1}}}^{2^n - 2^i} \frac{y'(s)}{(2^n - s)^\alpha} ds. \tag{10.9}$$

We now make the substitution $s = 2^n - 2^i u$ in the integrals in the sum to obtain

$$D^\alpha y(t)|_{t=2^n} = D_0 + \sum_{i=1}^{n-1} \frac{1}{2^{i\alpha}} \int_{\frac{1}{2}}^{1} \frac{y'(2^n - 2^i us)}{u^\alpha} du. \tag{10.10}$$

Clearly if we calculate the weights for one of the integrals in the sum, by the methods previously given, then the weights for the next integral will be the same weights scaled by a factor of $1/2^\alpha$ and the step length will be scaled by a factor of 2. Different scaling factors are possible if we chose a different fundamental interval.

In the next theorem we prove that combining the fractal sum interpretation with any reasonable numerical method for obtaining an approximate solution of a fractional differential equation does not introduce any problems.

Theorem 10.1 *Suppose we have a convolution method for approximating a fractional differential operator,*

$$D^\alpha y(2^m) = \sum_{i=0}^{2^m/h} w_{2^m/h-i} y_i + \mathcal{O}(h^p) = \bar{D}^\alpha y(2^m) + \mathcal{O}(h^p), \tag{10.11}$$

of order p as the step length $h \to 0$. Let $m \in \mathbb{N}$ and $2^m/h \in \mathbb{N}$ and y be a piecewise continuous function bounded on $[0, \infty)$, which has only a finite number of discontinuities on any finite interval. Then applying the fractal sum decomposition to the numerical method does not degrade the convergence or stability of the numerical method.

Proof Define the function $y_i^*(t)$ as follows

$$\begin{aligned} y_1^*(t) &= \begin{array}{l} y(t),\ t \in [0, 0.5] \\ y(t),\ t \in (0.5, 1] \end{array} \\ y_i^*(t) &= \begin{array}{l} 0,\ t \in [0, 0.5] \\ y(2^i t)/2^{i\alpha},\ t \in (0.5, 1] \end{array} \end{aligned}$$

for $i = 1, \ldots, m$. Then for each i, y_i^* is a piecewise continuous function bounded on $[0, 1]$ with only a finite number of discontinuities. Therefore y_i^* satisfies the conditions required for the numerical method to have order p as $h \to 0$. Now since

$$D^\alpha(y)|_{t=2^m} = \sum_{i=1}^{m} D^\alpha y_i^*|_{t=1}, \tag{10.12}$$

we must also have

$$\bar{D}^\alpha(y)|_{t=2^m} + \mathcal{O}(h^p) = \sum_{i=1}^{m} (\bar{D}^\alpha y_i^*|_{t=1} + \mathcal{O}(h^p)), \tag{10.13}$$

which is to say

$$\bar{D}^\alpha(y)|_{t=2^m} = \sum_{i=1}^{m} \bar{D}^\alpha y_i^*|_{t=1} + \mathcal{O}(h^p). \tag{10.14}$$

In the next section we implement the above idea as a simple algorithm (i.e. an algorithm that is simple to implement) which only partially meets the requirements of Theorem 10.1. In the final section we will work out an implementation which fully meets the requirements of Theorem 10.1.

10.3 Nested Mesh

In the Nested Mesh implementation our object is to minimise the computational effort and the size of the algorithm.

For convenience we take the most recent interval $[t-1, t]$ of the fractional derivative and choose a step length h such that $nh = 1$ for some $n \in \mathbb{N}$. To

$I_1 = [t-1, t]$	h
$I_2 = [t-2, t-1]$	$2h$
$I_3 = I[t-4, t-2]$	$4h$
$I_4 = [t-8, t-4]$	$8h$
...	...
$I_m = [t-2^m, t-2^{m-1}]$	$2^m h$
...	...

Table 10.1: Intervals and associated step lengths.

approximate the fractional derivative over $[t-1, t]$ we calculate the weights required to approximate

$$D_t = \frac{1}{\Gamma(1-\alpha)} \int_{t-1}^{t} \frac{y'(u)}{(t-u)^{\alpha}} du, \tag{10.15}$$

which are the same as for the fractional trapezium rule. Thus we calculate the set of weights $\omega_0, \ldots, \omega_n, \omega_{n+1}, \ldots, \omega_{2n}$ based on a step length of h.

Now let $\Omega_0 = (\omega_0, \ldots, \omega_n)$ and $\Omega_1 = (\omega_{n+1}, \ldots, \omega_{2n})$. From Ω_1 we generate the weight sets Ω_i such that $\Omega_i = \Omega_1 / 2^{\alpha(i-1)}$, $i = 1, \ldots, n-1$ together with an associate step length of $2^{i-1}h$, see Table 10.1. Thus the total set of convolution weights Ω can be generated by the recursive union

$$\Omega = \Omega_0 \cup \bigcup_{i=1}^{\infty} \frac{\Omega_1}{2^{\alpha(i-1)}}.$$

In this implementation we have introduced two potential sources of additional error. Firstly in the interval $[0, t-1)$ on iterations where $j \neq 0 \mod 2^m$ there will be an error because the weights are slightly offset and secondly, for the same condition on j, an initial interval of, at most, length $(2^m - 1)h$ will be truncated.

We can bound the error introduced by the weight offset by considering the difference δ between two successive weights in Ω_m for $m > 1$. Since

the weights, in each set, are decreasing we need only consider the first two weights in Ω_m. Thus we have

$$\delta = \frac{1}{2^{(m-1)\alpha}}(\omega_{n+1,j} - \omega_{n+2,j}).$$

For the fractional trapezium rule, expanding by the binomial theorem, gives a maximum error through weight offset of $\mathcal{O}(j^{-\alpha}/2^{(m-1)\alpha})$, since n is constant. Thus the error in the convolution sum caused by weight offset error E_w can be bounded by

$$E_w = y_{max} 2^{m\alpha} hn\mathcal{O}(j^{-\alpha}) = \mathcal{O}(j^{-1}), \tag{10.16}$$

since $m = \mathcal{O}(\log_2 j)$. Combining this with Theorem 6.1 we find we have reduced the order of the method to $\mathcal{O}(j^{-1})$.

10.3.1 Computational Assays

In Figure 10.1 we have plotted the approximate solution to the system (10.17),

$$D^{0.5}y = -y + t + t^{0.5}/\Gamma(1.5),\ y_0 = 0, \tag{10.17}$$

using the fractional trapezium rule in the Nested Mesh algorithm, with an initial window of length 1 and a step length $h = 1/4$. The upper line is the exact solution and the lower line is its approximation. Even under these crude conditions we see that the Nested Mesh algorithm gives a much better approximation than the finite memory principle. In Figure 10.2 the lower curve is the error in the approximation, the upper curve is a plot of $\log t$.

We have calculated the approximate solution to the system

$$D^{0.5}y = -y,\ y_0 = 1, \tag{10.18}$$

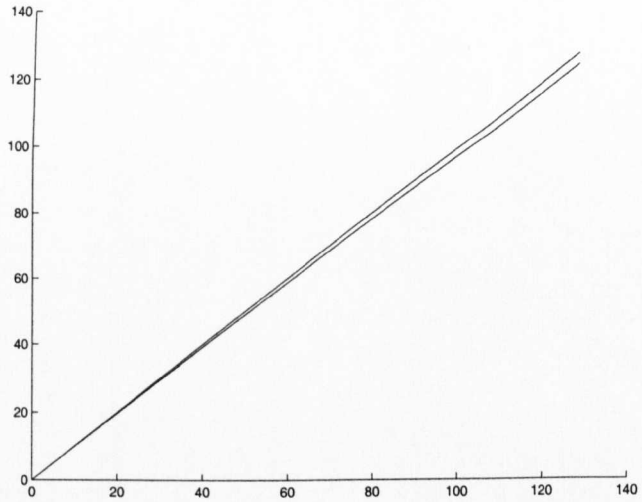

Figure 10.1: Approximate solution to system (10.17).

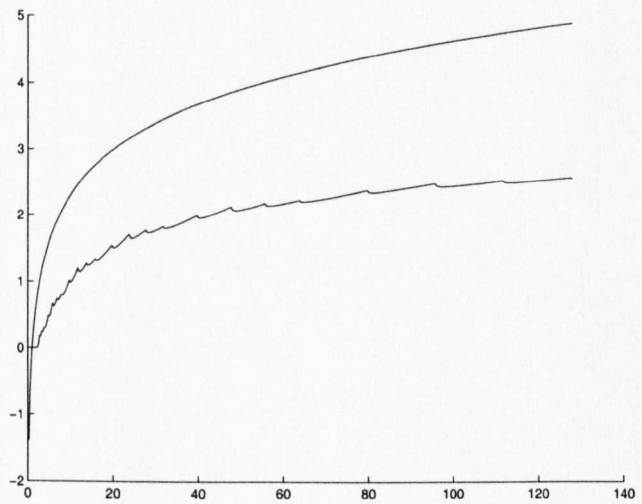

Figure 10.2: Error in the approximate solution to system (10.17).

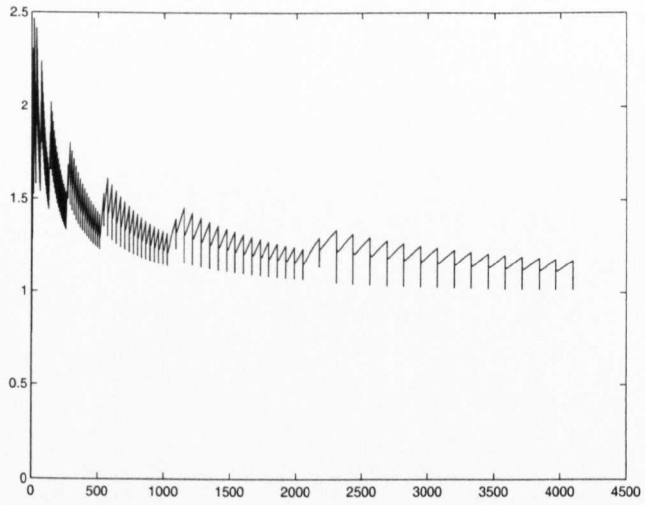

Figure 10.3: Percentage divergence in the approximate solution to system (10.18) for $h = 1/16$.

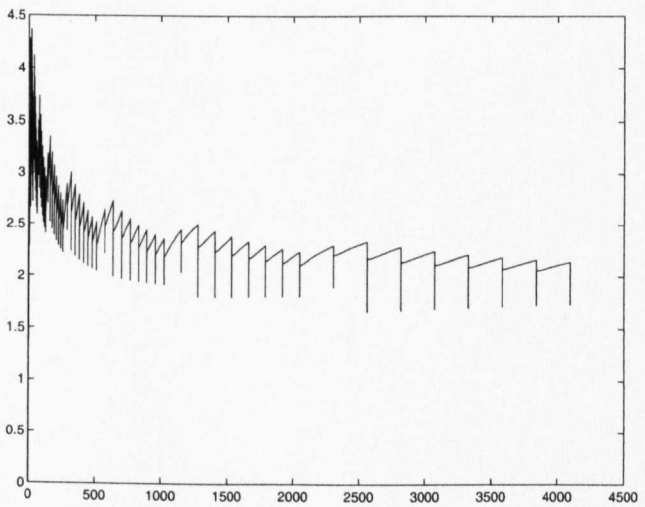

Figure 10.4: Percentage divergence in the approximate solution to system (10.18) for $h = 1/8$.

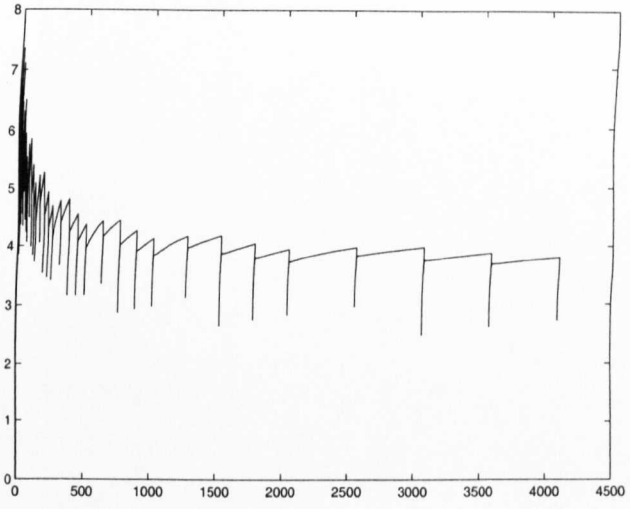

Figure 10.5: Percentage divergence in the approximate solution to system (10.18) for $h = 1/4$.

produced by using the unmodified fractional trapezium rule and the modified fractional trapezium rule using an initial window of length 1.

10.3.2 Effects of Scaling

To see the effects of using different scaling factors we have repeated calculations of the type given above for system (10.18) using scaling factors of 5, where we reuse 80% of the weights and expand the step length by a factor of 5, and 10 where we reuse 90% of the weights and expand the step length by a factor of 10, as well as using a scaling factor of 2. We summarise the computational effort in table 10.2.

In Figures 10.3, 10.4 and 10.5 we have plotted the percentage difference between the approximations for step lengths of 2^{-4}, 2^{-3} and 2^{-2} to the system (10.18). As we can see as the step length is halved the difference between the two approximations reduces $1/\sqrt{2}$. This is as we would anticipate from

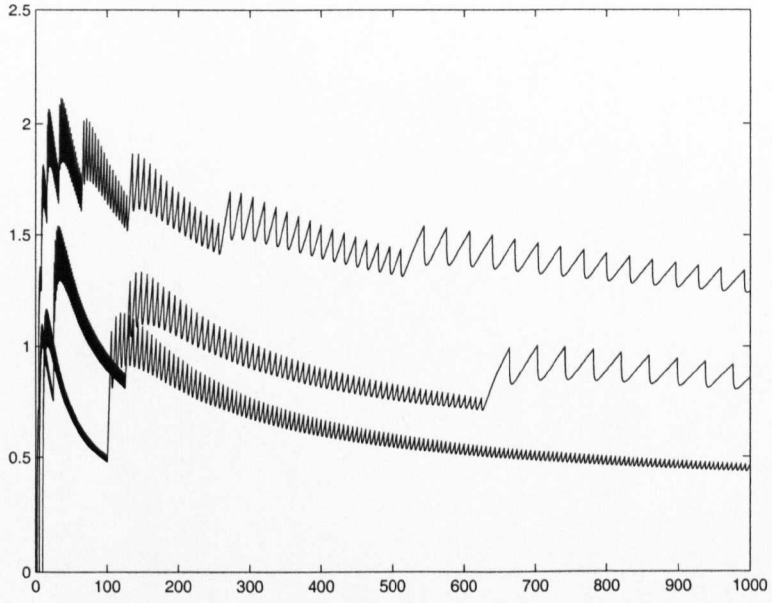

Figure 10.6: Percentage difference in the approximation to system (10.18) for three different scaling factors over the interval $[0, 1000]$ with a step length of 2^{-4}.

the reduction in order due to weights being *off grid*, however in the limit the approximation calculated using the Nested Mesh algorithm will converge to the approximation calculated using the original algorithm.

In Figures 10.6, 10.7 and 10.8 we have plotted the difference between the approximate solution to (10.18) with full integration and nested mesh integration for three scaling factors using a step length of 2^{-4}.

In Figures 10.9 and 10.10 we repeat the above calculations using a larger step length of 2^{-2} and observe the expected consistency.

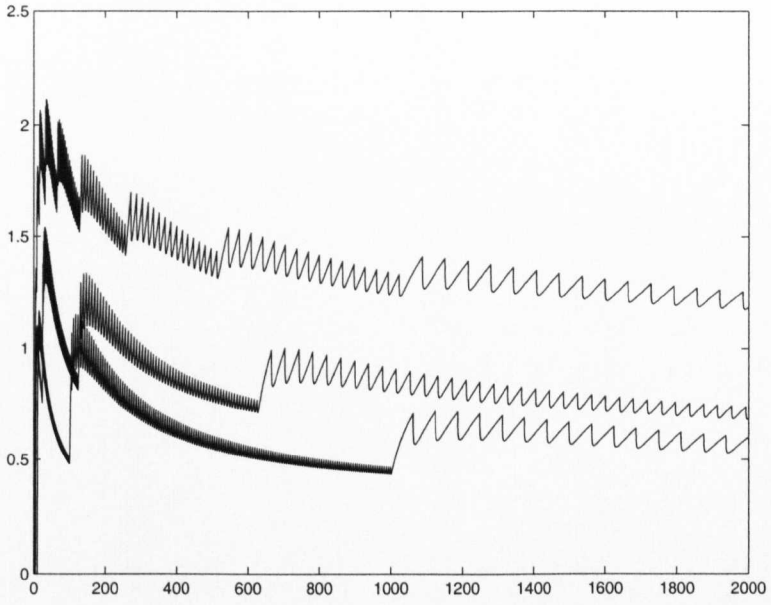

Figure 10.7: Percentage difference in the approximation to system (10.18) for three different scaling factors over the interval $[0, 2000]$ with a step length of 2^{-4}.

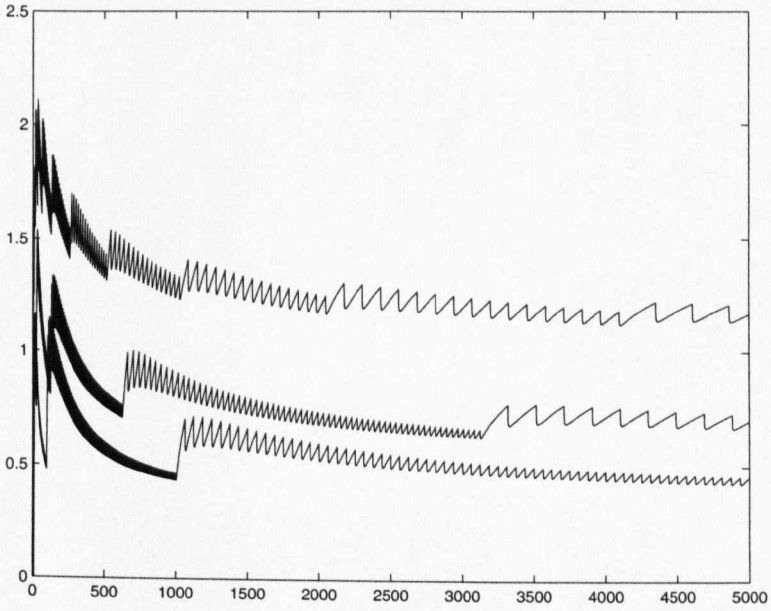

Figure 10.8: Percentage difference in the approximation to system (10.18) for three different scaling factors over the interval $[0, 5000]$ with a step length of 2^{-4}.

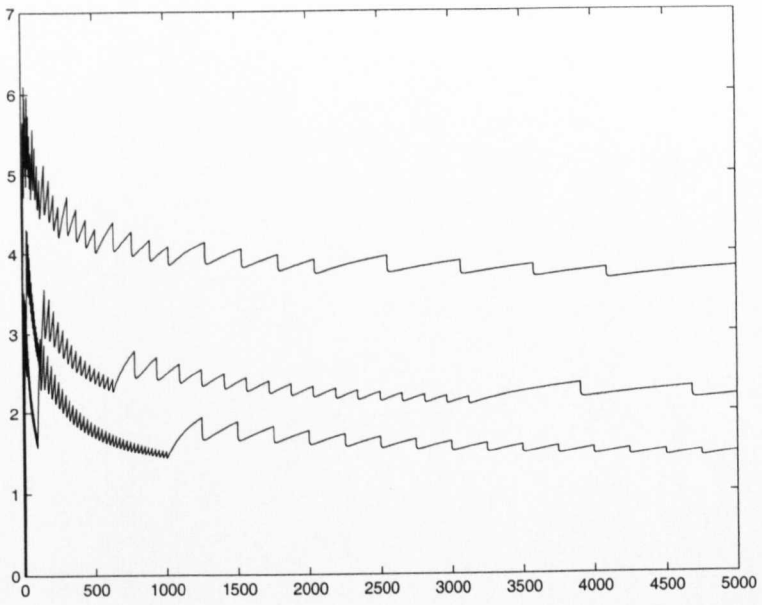

Figure 10.9: Percentage difference in the approximation to system (10.18) for three different scaling factors over the interval $[0, 5000]$ with a step length of 2^{-2}.

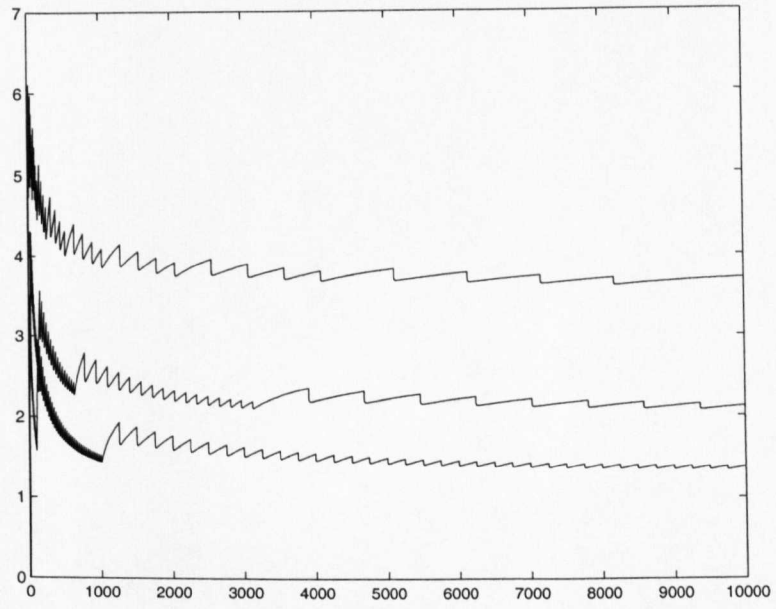

Figure 10.10: Percentage difference in the approximation to system (10.18) for three different scaling factors over the interval $[0, 10000]$ with a step length of 2^{-2}.

Scale Factor	125	250	500	1000	2000	5000
2	0.66	1.51	3.40	7.57	16.67	46.76
5	1.00	2.34	5.29	12.01	26.43	75.36
10	1.45	3.40	7.65	17.59	39.77	110.93
∞	4.05	16.09	64.18	256.37	1024.74	6401.84

Table 10.2: Time and computational effort for a variety of dilation scaling factors with a step length of 2^{-4}.

10.4 Fractal Sum

In the Nested Mesh implementation errors are introduced by applying the *iterated* weights *off grid* and by a small truncation. In each of theses cases the error is under the control of the step length and vanishes as $h \to 0$. However as we have seen the weights being off grid affects the order of convergence. In this section we will show how to calculate the *on grid* weights to implement the fractional trapezium rule without any truncation.

Recollect we are approximating a fractional differential equation using a step length of h. Now because of the scaling property of the fractional differential operator, using the sum of integrals over the same interval formulation, we can interpret the problem as using the same step length in each of the integrals since they are all over the same interval, then the convergence of our modification follows from that of the unmodified method. However to solve for an accurate approximation of the unknown function, in the fractional differential equation, we need to calculate the weights with respect to their position on the h-mesh throughout the interval of integration.

Therefore to eliminate the two sources of error we had in the Nested Mesh implementation we need the exact weights for each D_i, with concomitant step

length $2^{i-1}h$, for each of the cases $j = 0 \mod 2^{i-1}$, $j = 1 \mod 2^{i-1}$, ..., $j = 2^{i-1} - 1 \mod 2^{i-1}$.

Firstly we consider the weights required for D_i when the whole of its interval of integration is included in $[0, t]$, i.e. when $t - 2^{i-1}T \geq 0$. We then consider the weights required for D_i when only part of its potential interval of integration is included in the interval covered by the fractional differential operator.

10.4.1 Calculation of the Weights

Let the weights calculated for D_{i-1} be Ω_{i-1} then for $j = 0 \mod 2$ we can obtain the weights for Ω_i by the mapping $\omega_j = \omega_{j/2}/2^\alpha$, however when $j > 2^{i-1}$ and $j = 1 \mod 2$ we will need to calculate some new weights. These are of three kinds: weights at the beginning of the interval, weights in the middle of the interval, and weights at the end of the interval.

Weights at the Beginning of the Interval

To deal with this case the interpolating function, at this point, takes the form

$$\begin{array}{ll} p_k(t) = \frac{t-(k-2^{i-1})h}{2^{i-1}h} & \text{for } t \in [(k-2^{i-1})h, kh), \\ p_k(t) = 1 & \text{for } t = kh. \end{array} \tag{10.19}$$

In the actual calculation k will be of the form $k = j - 2^{i-1}n$, where $j > 2^{i-1}n$.

Integrating, as before, to obtain the weight we have

$$I_k = \frac{1}{2^{i-1}h} \int_{(k-2^{i-1})h}^{kh} \frac{s - (k-2^{i-1})h}{(jh-s)^{1+\alpha}} ds.$$

Making the substitution $s = (j-u)h$ gives

$$\begin{aligned} I_k &= \frac{1}{2^{i-1}h^\alpha}\int_{j-(k-2^{i-1})}^{j-k} \frac{u-(j-(k-2^{i-1}))}{u^{1+\alpha}}du, \\ &= \frac{1}{2^{i-1}h^\alpha}[\frac{u^{1-\alpha}}{1-\alpha} + \frac{(j-(k-2^{i-1}))u^{-\alpha}}{\alpha}]_{j-(k-2^{i-1})}^{j-k}, \end{aligned}$$

which gives

$$\alpha(1-\alpha)2^{i-1}h^\alpha I_k = (j-k)^{1-\alpha} - (j-(k-2^{i-1}))^{1-\alpha} + 2^{i-1}(1-\alpha)(j-k)^{-\alpha}.$$

Weights in the Middle of the Interval

To calculate the new middle weights required for Ω_i we use the interpolating functions

$$\begin{array}{lll} p_k(t) = \frac{t-(k-2^{i-1})h}{2^{i-1}h} & \text{for} & t \in [(k-2^{i-1})h, kh), \\ p_k(t) = 1 & \text{for} & t = kh, \\ p_k(t) = \frac{(k+2^{i-1})h-t}{2^{i-1}h} & \text{for} & t \in (kh, (k+2^{i-1})]. \end{array} \tag{10.20}$$

Integrating the interpolating functions gives

$$I_k = \frac{1}{2^{i-1}h}\int_{k-2^{i-1}h}^{kh} \frac{s-(k-2^{i-1})h}{(jh-s)^{1+\alpha}}ds + \frac{1}{2^{i-1}h}\int_{kh}^{k+2^{i-1}h} \frac{(k+2^{i-1})h-s}{(jh-s)^{1+\alpha}}ds. \tag{10.21}$$

For $k = j$ and k odd we must carry out the integration to obtain the required weights. Making the substitution $s = (j-u)h$ as before we have

$$\begin{aligned} 2^{i-1}h^\alpha I_k &= -\int_{j-(k-2^{i-1})}^{j-k} \frac{(j-(k-2^{i-1}))-u}{u^{1+\alpha}}du - \int_{j-k}^{j-(k+2^{i-1})} \frac{u-(j-(k+2^{i-1}))}{u^{1+\alpha}}du \\ &= \int_{j-k}^{j-(k-2^{i-1})} \frac{(j-(k-2^{i-1}))-u}{u^{1+\alpha}}du - \int_{j-(k+2^{i-1})}^{j-k} \frac{(j-(k+2^{i-1}))-u}{u^{1+\alpha}}du \\ &= [-\frac{(j-(k-2^{i-1}))}{\alpha u^\alpha} - \frac{u^{1-\alpha}}{1-\alpha}]_{j-k}^{j-(k-2^{i-1})} - [-\frac{(j-(k+2^{i-1}))}{\alpha u^\alpha} - \frac{-u^{1-\alpha}}{1-\alpha}]_{j-(k+2^{i-1})}^{j-k} \end{aligned}$$

which gives

$$\begin{aligned} 2^{i-1}h^{\alpha}I_k &= (-\frac{(j-(k-2^{i-1}))^{1-\alpha}}{\alpha} - \frac{(j-(k-2^{i-1}))^{1-\alpha}}{1-\alpha} \\ &+ \frac{(j-(k-2^{i-1}))}{\alpha(j-k)^{\alpha}} + \frac{(j-k))^{1-\alpha}}{1-\alpha} \\ &+ \frac{(j-(k+2^{i-1}))}{\alpha(j-k)^{\alpha}} + \frac{(j-k)^{1-\alpha}}{1-\alpha} \\ &- \frac{(j-(k+2^{i-1}))^{(1-\alpha)}}{\alpha} - \frac{(j-(k+2^{i-1}))^{1-\alpha}}{1-\alpha}). \end{aligned}$$

Hence

$$2^{i-1}\alpha(1-\alpha)h^{\alpha}I_k = 2(j-k))^{1-\alpha} - (j-(k-2^{i-1}))^{1-\alpha} - (j-(k+2^{i-1}))^{(1-\alpha)}.$$

We see that for $k = 2k' \in \mathbb{N}$ and j even we have $I_{k'} = I_k/2^{\alpha}$ as we would expect.

Weights at the End of the Step Length $= 2^{i-1}h$ Region

To deal with this case the interpolating function, at this point, takes the form

$$\begin{aligned} p_k(t) &= 1 \qquad\qquad for \quad t = kh. \\ p_k(t) &= \tfrac{(k+2^{i-1})h-t}{2^{i-1}h} \quad for \quad t \in (kh, (k+2^{i-1})h]. \end{aligned} \tag{10.22}$$

Integrating, as before, to obtain the weight we have

$$I_k = \frac{1}{2^{i-1}h}\int_{kh}^{(k+2^{i-1})h} \frac{(k+2^{i-1})h - s}{(jh-s)^{1+\alpha}} ds.$$

Making the substitution $s = (j-u)h$ gives

$$\begin{aligned} I_k &= \frac{1}{2^{i-1}h^{\alpha}}\int_{j-(k+2^{i-1})}^{j-k} \frac{u-(j-(k+2^{i-1}))}{u^{1+\alpha}} du, \\ &= \frac{1}{2^{i-1}h^{\alpha}}[\frac{u^{1-\alpha}}{1-\alpha} + \frac{(j-(k+2^{i-1}))u^{-\alpha}}{\alpha} -]_{j-(k+2^{i-1})}^{j-k}, \end{aligned}$$

which gives

$$\alpha(1-\alpha)2^{i-1}h^{\alpha}I_k = (j-k)^{1-\alpha} - (j-(k+2^{i-1}))^{1-\alpha} - 2^{i-1}(1-\alpha)(j-k)^{-\alpha}.$$

Weights For the End of the Interval

We need to calculate two weights to eliminate the truncation, which occurred in the Nested Mesh implementation. When $j \neq 0 \mod 2^{i-1}$. Let $0 < d < 2^{i-1}$ then firstly we need

$$I_{d_1} = \int_0^{dh} \frac{s}{dh(jh-s)^{1+\alpha}} ds, \tag{10.23}$$

making the substitution $s = jh - uh$ gives

$$\begin{aligned} I_{d_1} &= \frac{1}{dh}\int_j^{j-d} \frac{(j-u)h}{(uh)^{1+\alpha}}(-hdu) \\ &= \frac{h^{-\alpha}}{d}\int_j^{j-d} \frac{(j-d)-u}{u^{1+\alpha}} du \\ &= \frac{h^{-\alpha}}{d}[\frac{u^{1-\alpha}}{1-\alpha} + \frac{ju^{-\alpha}}{\alpha}]_j^{j-d} \\ &= \frac{h^{-\alpha}}{d}[\frac{(j-d)^{1-\alpha}}{1-\alpha} + \frac{j(j-d)^{-\alpha}}{\alpha} - \frac{j^{1-\alpha}}{1-\alpha} - \frac{j^{1-\alpha}}{\alpha}] \end{aligned}$$

which gives

$$d\alpha(1-\alpha)h^{\alpha} I_{d_1} = (j-d)^{1-\alpha} - j^{1-\alpha} + d(1-\alpha)(j-d)^{-\alpha}. \tag{10.24}$$

Secondly we need

$$I_{d_2} = \int_0^{dh} \frac{dh-s}{dh(jh-s)^{1+\alpha}} ds, \tag{10.25}$$

making the substitution $s = jh - uh$ gives

$$\begin{aligned} I_{d_2} &= \frac{1}{dh}\int_j^{j-d} \frac{((j-d)-u)h}{(uh)^{1+\alpha}}(hdu) \\ &= \frac{h^{-\alpha}}{d}\int_j^{j-d} \frac{u-(j-d)}{u^{1+\alpha}} du \\ &= \frac{h^{-\alpha}}{d}[\frac{u^{1-\alpha}}{1-\alpha} + \frac{(j-d)u^{-\alpha}}{\alpha}]_j^{j-d} \\ &= \frac{h^{-\alpha}}{d}[\frac{(j-d)^{1-\alpha}}{1-\alpha} + \frac{(j-d)^{1-\alpha}}{\alpha} - \frac{j^{1-\alpha}}{1-\alpha} - \frac{(j-d)j^{-\alpha}}{\alpha}] \end{aligned}$$

which gives

$$d\alpha(1-\alpha)h^{\alpha}I_{d_2} = (j-d)^{1-\alpha} - j^{1-\alpha} + d(1-\alpha)j^{-\alpha}. \tag{10.26}$$

Chapter 11

Multiterm Equations and FLMM

11.1 Motivation

In this chapter we discuss methods for approximating the solution of linear multiterm fractional differential equations of the form

$$\sum_{j=0}^{p} c_j D^{\beta_j} y = f, \tag{11.1}$$

where $0 \leq \beta_0 < \cdots < \beta_r < \beta_{r+1} < \cdots < \beta_p$, $c_s \in \mathbb{R}$ and $c_p = 1$, which arise out of the use of Laplace transform methods to obtain its analytical solution.

There are two approaches we can take to approximating the solution of (11.1). Firstly we can take the one operator approach and collect each of the fractional differential operators into one integral operator and attempt to approximate its action. This is the approach that we discuss in this chapter. Alternatively we can treat the fractional differential equation as being built up by the application of many fractional differential operators and treat it as a system of fractional differential equations each of order at most one, we discuss this approach in chapter 12.

11.2 Multiterm Equations and FLMM

In [25] a method for obtaining the analytic solution of a multiterm fractional differential equation of the form of equation (11.1) is developed which is essentially based on using Laplace transform theory. The analytical solution to equation (11.1) was given in Theorem 3.12.

We recollect that the CQM, given in Chapter 7 is also based on Laplace transform theory. In [16] the feasibility of using CQM to give a method for approximating the solution of a multiterm fractional differential equation of the form of equation (11.1) is investigated. The following theorem is given which shows that CQM maintain their convergence properties in this case.

Theorem 11.1 *([16] 7)*

Let E_h, obtained by a CQM which satisfies the conditions of chapter 7, be the approximate to multivariate Mittag-Leffler function $E(t)$ which solves the homogeneous version of the initial value problem given in (11.1) then at $t_n = nh$ we have

$$|E_h(nh) - E(t_n)| = \mathcal{O}(h^p),\ h \to 0. \tag{11.2}$$

An equivalent theorem is given for some cases of the inhomogeneous problem.

Theorem 11.2 *([16] 9)*

Let E_h, obtained by a CQM which satisfies the conditions of chapter 7, be the approximate inhomogeneous version of the initial value problem given in (11.1), where $f(t) = t^{\gamma-1}\tilde{f}(t)$ for some $\tilde{f} \in C^p[0, t_n]$, and let $E(t)$ be the true solution to (11.1), then at $t_n = nh$ we have

$$|E_h(nh) - E(t_n)| = \mathcal{O}(h^q),\ h \to 0,\ q = \min\{\gamma, p\}. \tag{11.3}$$

In [16] extensive numerical examples are given which demonstrate that the proven convergence does occur, at least for BDF2. The example equations given contain only one fractional differential term so the kernel of the integral equation, whose Laplace transform supplies the weights, is of a simple form. No indication is give of the amount of computational effort involved.

From equation (9) in [16] we conclude that the Laplace transform of the kernel $\bar{K}(s)$ of equation (11.1) will take the form

$$\bar{K}(s) = \frac{s^{\beta_p - m_p}}{\sum_{j=0}^{p} c_j s^{\beta_j}}. \tag{11.4}$$

Therefore we will need to find the formal power series expansion of the ratio of the characteristic polynomials of the LMM for each β_j, $j = 0, \ldots, p$, and for $\beta_p - m_p$, sum these expansions for $j = 0, \ldots, p$, find the reciprocal of this sum and multiply it by the expansion for $\beta_p - m_p$. As we saw in the examples given in chapter 7 calculating just one expansion requires more computational effort that actually calculating the sum of the convolution integral.

Whilst this method undoubtedly has excellent theoretical properties it may prove to be computationally inaccessible.

Chapter 12

Systems of Equations

12.1 Motivation

Miller and Ross [45] give a method for calculating the analytic solution to multiterm fractional differential equations of the form

$$[D^{n\nu} + a_1 D^{(n-1)\nu} + \cdots + a_{n-1} D^{\nu} + a_n D^0] y(t) = 0,\ \nu = \frac{1}{q},\ q \in \mathbb{N}. \quad (12.1)$$

The fractional operator $D^{i\nu}$ is interpreted as the Riemann-Liouville fractional derivative of order ν applied i-times. Under this interpretation they describe the multiterm fractional differential equation as a sequential fractional differential equation.

Ford and Diethelm [22] have shown that this interpretation arises naturally when using the Caputo fractional derivative, and subsequently developed an algorithm for approximating the solution of a multiterm fractional differential equation by treating it as a system of equations of fractional differential order ν.

Given the density of the rational numbers any fractional differential equation can be modelled arbitrarily well by this method, the finite precision arithmetic of computers being a considerably more limiting constraint on

accuracy than any number theoretic compromises. This approach, however, may result in a very large system of equations.

12.2 Systems of Equations

In this section we develop a numerical method for solving fractional differential equations based on treating the fractional differential equation as a mixed system of equations of orders in the interval $(0, 1]$. This produces the system of the lowest possible dimension under the constraint of the maximum permissible order being 1. This can be seen as an extension of the method used for solving high order differential equations whilst avoiding the problems of order blow up that can potentially occur with the method in [22].

Discretisation of Derivatives

For the sake of clarity and simplicity we will discretise integer order differential equations using the trapezium rule:

$$Dy = f \Rightarrow y_i = y_{i-1} + \frac{1}{2}h(f_i + f_{i-1}),$$

and discretise the fractional differential equations using the fractional trapezium rule:

$$D^\alpha y = \frac{1}{\gamma_i}(\sum_{k=0}^{i} \omega_{k,i} y_{i-k} + \frac{y_0}{\alpha}), \tag{12.2}$$

where $\gamma_i = (ih)^\alpha \Gamma(-\alpha)$.

It should be noted that D and D^α are operators, their appearance in matrices is a notational device, therefore questions about the singularity of matrices do not arise until after the derivatives have been discretised.

12.2.1 Systems of Equations for Ordinary Differential Equations

Recollect that to solve the equation

$$D^n y + \cdots + b_1 y = f,$$
$$y^{(i)}(0) = y_0^{(i)},\ i = 0, \ldots, n-1,$$

as a system of first order equations we let

$$\begin{aligned} {}^1Y &= y \\ {}^{i+1}Y &= D^i y,\ i = 1, \ldots, n. \end{aligned}$$

Expressed in matrix notation this gives

$$\begin{pmatrix} D & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & D \end{pmatrix} \begin{pmatrix} {}^1Y \\ \vdots \\ {}^nY \end{pmatrix} = \begin{pmatrix} {}^2Y \\ \vdots \\ -\sum_{i=1}^n b_i\, {}^iY + f \end{pmatrix}.$$

To obtain an approximate solution to the ordinary differential equations we discretise the derivatives and then solve the resulting matrix system to obtain the solution y.

12.2.2 Linear Fractional Differential Equations

Ordering of Fractional Derivatives

Consider the fractional differential equation

$$Dy + a_1 D^{\alpha_1} y + a_2 D^{\alpha_2} + a_3 y = f,\ 1 > \alpha_1 > \alpha_2 > 0. \tag{12.3}$$

Our aim is to write the discrete approximation to (12.3) in the form

$$A_i \mathbf{Y}_i = \mathbf{H}_i,$$

where $\mathbf{H}_i$ is composed of combinations of elements of $\mathbf{Y}_0, \ldots, \mathbf{Y}_{i-1}$ and $f_0, \ldots, f_{i-1}$, f_i, in such a way that the dimension of $\mathbf{Y}_i$ is minimal.

There are two possible ways we can write this as a system where each system has the same dimension. Firstly we can write the fractional differential operator as acting on the first component of $\mathbf{Y}$ which gives:

$$\begin{aligned} {}^1Y &= y, \\ {}^2Y &= D^{\alpha_2}\,{}^1Y, \\ {}^3Y &= D^{\alpha_1}\,{}^1Y, \\ {}^4Y &= D\,{}^1Y, \end{aligned}$$

secondly we could write the system as a sequence of fractional derivatives each acting on successive components of $\mathbf{Y}$:

$$\begin{aligned} {}^1Y &= y, \\ {}^2Y &= D^{\alpha_2}\,{}^1Y, \\ {}^3Y &= D^{\alpha_1-\alpha_2}\,{}^2Y, \\ {}^4Y &= D^{1-\alpha_1}\,{}^3Y. \end{aligned}$$

We choose to develop the first alternative since it requires the weights for one less fractional derivative to be calculated (and therefore one less convolution sum per iteration) and has less propagation of rounding error.

12.2.3 The Five Term Second Order Fractional Differential Equation

We start by giving methods to obtain approximate numerical solutions to multiterm fractional differential equations which have already appeared in

the scientific literature. We then give a method sufficient for any multiterm fractional differential equation.

By the five term second order fractional differential equation we mean an equation of the form

$$D^2y + b_3D^{1+\beta}y + b_2Dy + b_2D^{\alpha}y + b_0y = f, \tag{12.4}$$
$$y(0) = y_0,\ y'(0) = y_0',\ \alpha, \beta \in (0,1),\ b_0 \neq 0.$$

Equation (12.4) includes, with some redundancy, the Bagley-Torvik equation as a special case.

We write equation (12.4) as the system

$$\begin{aligned} {}^1Y &= y, \\ {}^2Y &= D^{\alpha}y, \\ {}^3Y &= Dy, \\ {}^4Y &= D^{\beta}Dy, \\ {}^5Y &= D^2y. \end{aligned}$$

In matrix form this gives

$$\begin{pmatrix} D^{\alpha} & 0 & 0 & 0 \\ D & 0 & 0 & 0 \\ 0 & 0 & D^{\beta} & 0 \\ 0 & 0 & D & 0 \end{pmatrix} \begin{pmatrix} {}^1Y \\ {}^2Y \\ {}^3Y \\ {}^4Y \end{pmatrix} = \begin{pmatrix} {}^2Y \\ {}^3Y \\ {}^4Y \\ f - \sum_{j=1}^{4} b_{j-1}\,{}^jY \end{pmatrix}.$$

Discretising and collecting terms together gives:

$$\begin{pmatrix} -{}^{\alpha}\omega_{0,i} & {}^{\alpha}\gamma_i & 0 & 0 \\ 1 & 0 & -\frac{h}{2} & 0 \\ 0 & 0 & -{}^{\beta}\omega_{o,i} & {}^{\beta}\gamma_i \\ \frac{hb_0}{2} & \frac{hb_1}{2} & 1+\frac{hb_2}{2} & \frac{hb_3}{2} \end{pmatrix} \begin{pmatrix} {}^1Y_i \\ {}^2Y_i \\ {}^3Y_i \\ {}^4Y_i \end{pmatrix} = \begin{pmatrix} {}^1S_i \\ {}^2S_i \\ {}^3S_i \\ {}^4S_i \end{pmatrix},$$

where

$$
\begin{aligned}
{}^1S_i &= \sum_{k=1}^{i} {}^{\alpha}\omega_{k,i}\, {}^1Y_{i-k} + \frac{{}^1Y_0}{\alpha}, \\
{}^2S_i &= {}^1Y_{i-1} + \frac{h}{2}\, {}^3Y_{i-1}, \\
{}^3S_i &= \sum_{k=1}^{i} {}^{\beta}\omega_{k,i}\, {}^3Y_{i-k} + \frac{{}^1Y_0}{\beta}, \\
{}^4S_i &= {}^3Y_{i-1} + F_i - \frac{h}{2}\sum_{j=1}^{4} b_{j-1}\, {}^jY_{i-1}.
\end{aligned}
$$

and $F_i = h(f_i + f_{i-1})/2$.

With initial conditions $y(0) = 1$, $y'(0) = 0$, we have calculated the solutions to the equations:

$$D^2y + D^{0.5}y + y = 0, \tag{12.5}$$

$$D^2y + Dy + y = 0, \tag{12.6}$$

$$D^2y + D^{1.5}y + y = 0, \tag{12.7}$$

in Figures 12.1, 12.2, and 12.3 respectively.

This shows the variety of system behaviour fractional generalisations of velocity damping may be used to model.

Theorem 12.1 *The order of convergence of the resulting method is equal to the lowest order of the methods that it comprises.*

Proof We prove this by induction. Clearly this is the case for the first iteration $\mathbf{Y}_1$ since we are using exact function values. Now let the order of the lowest order method be p then by hypothesis at the n-th iteration all the components of $\mathbf{Y}_n$ are of $\mathcal{O}(h^p)$ i.e. $y_n = {}^1Y_n + \mathcal{O}(h^p)$, therefore at the

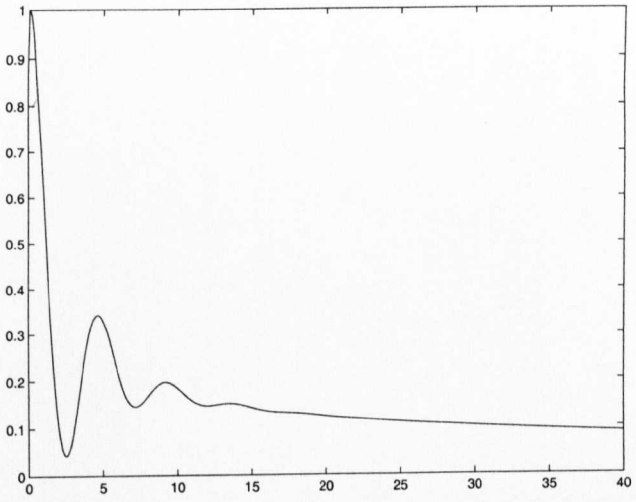

Figure 12.1: Approximate solution to equation (12.5).

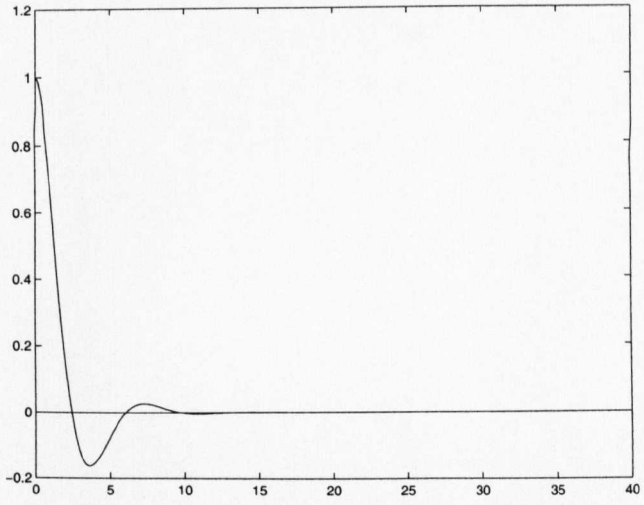

Figure 12.2: Approximate solution to equation (12.6).

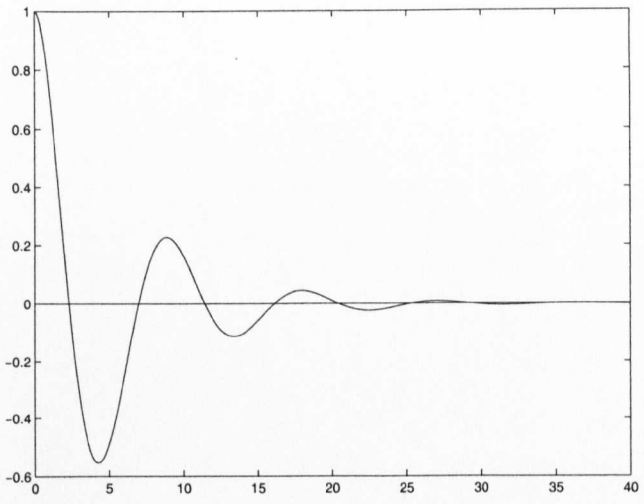

Figure 12.3: Approximate solution to equation (12.7).

$n+1$-iteration $\mathbf{Y}_{n+1}$ will include the accumulated order $\mathcal{O}(h^p)$ from the n previous iterations plus the $\mathcal{O}(h^p)$ error introduced at the $n+1$-th iteration the combination of which is $\mathcal{O}(h^p)$.

12.2.4 Three Equations of Special Importance

We now show how our method can be applied to three fractional differential equations which have found practical application.

The Bagley-Torvik Equation

As before by the Bagley-Torvik equation we mean any linear fractional differential equation (12.8) of the form:

$$D^2y + b_2D^{1+\alpha}y + b_1y = f, \tag{12.8}$$

$$y(0) = y_0,\ y'(0) = y_0',\ \alpha \in (0,1),\ b_1b_2 \neq 0.$$

By the linearity of D and D^α we can rewrite this as

$$\begin{aligned} {}^1Y &= y, \\ {}^2Y &= Dy, \\ {}^3Y &= D^\alpha Dy, \\ {}^4Y &= D^2 y. \end{aligned}$$

In matrix notation this is

$$\begin{pmatrix} D & 0 & 0 \\ 0 & D^\alpha & 0 \\ 0 & D & 0 \end{pmatrix} \begin{pmatrix} {}^1Y \\ {}^2Y \\ {}^3Y \end{pmatrix} = \begin{pmatrix} {}^2Y \\ {}^3Y \\ f - b_2\,{}^3Y - b_1\,{}^1Y \end{pmatrix}.$$

Discretising as above this results in the matrix system

$$\begin{pmatrix} 1 & -\frac{h}{2} & 0 \\ 0 & -\omega_{0,i} & \gamma_i \\ \tau_1 & 1 & \tau_2 \end{pmatrix} \begin{pmatrix} {}^1Y_i \\ {}^2Y_i \\ {}^3Y_i \end{pmatrix} =$$

$$\begin{pmatrix} {}^1Y_{i-1} + \frac{h}{2}\,{}^2Y_{i-1} \\ {}^2S_i \\ F_i - \tau_1\,{}^1Y_{i-1} + {}^2Y_{i-1} - \tau_2\,{}^3Y_{i-1} \end{pmatrix},$$

where $F_i = h(f_i + f_{i-1})/2$, $\tau_1 = hb_1/2$, and $\tau_2 = hb_2/2$.

We have calculated the solutions in Figures 12.4, 12.5, and 12.6 to the example Bagley-Torvik equation:

$$D^2 y + D^{1+\alpha} y + y = 0,\ y(0) = 1,\ y'(0) = 0,$$

for $\alpha = 0.25,\ 0.5,\ 0.75$ respectively.

Fractional Oscillation Equation

We now apply the above techniques to the fractional oscillation equation

$$D^{1+\alpha} y(t) + by(t) = f(t),\ y(0) = y_0,\ y^{(1)}(0) = y_0^{(1)}, \tag{12.9}$$

where $\alpha \in (0,1)$, $t \in [0,\infty)$, and f satisfies a Lipschitz condition in t.

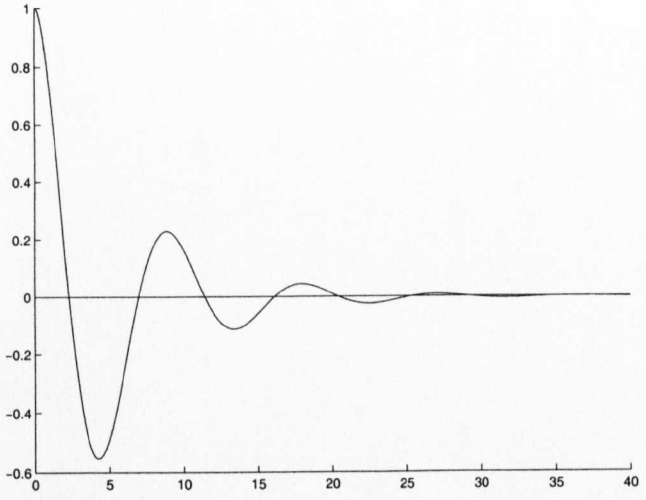

Figure 12.4: Approximate solution to equation (12.9) for $\alpha = 0.5$.

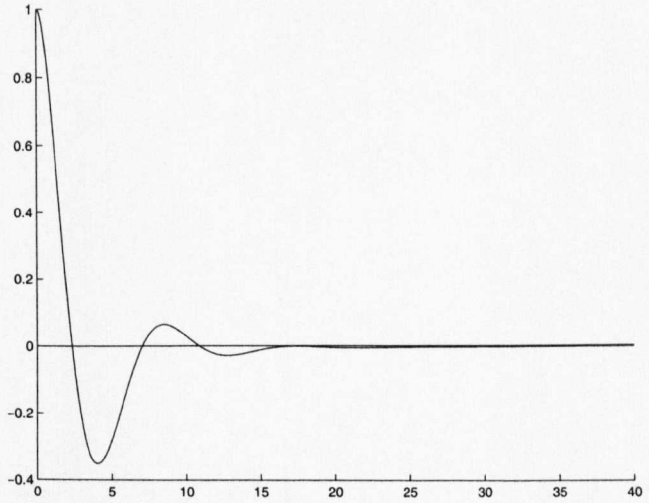

Figure 12.5: Approximate solution to equation (12.9) for $\alpha = 0.25$.

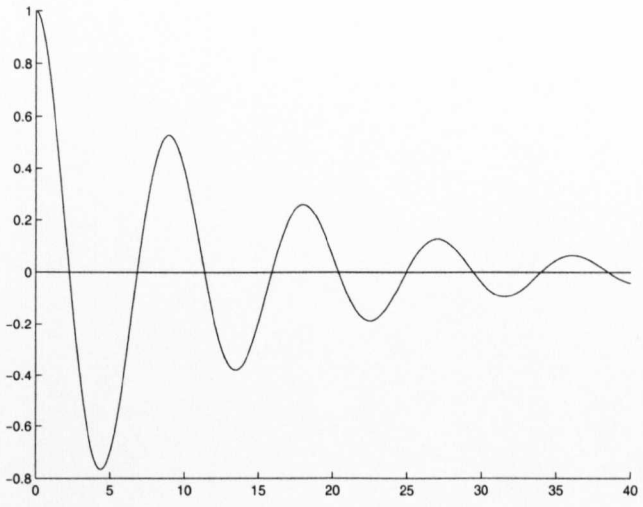

Figure 12.6: Approximate solution to equation (12.9) for $\alpha = 0.75$.

By the linearity of D and D^α we have

$$\begin{aligned} {}^1Y &= y, \\ {}^2Y &= Dy, \\ {}^3Y &= D^\alpha Dy. \end{aligned}$$

In matrix form this gives

$$\begin{pmatrix} D & 0 \\ 0 & D^\alpha \end{pmatrix} \begin{pmatrix} {}^1Y \\ {}^2Y \end{pmatrix} = \begin{pmatrix} {}^2Y \\ -b\,{}^1Y + f(t) \end{pmatrix}.$$

Discretising the derivatives as previously we have

$${}^1Y_i = {}^1Y_{i-1} + ({}^2Y_i + {}^2Y_{i-1})h/2,$$

and for the second equation we have

$$-b\,{}^1Y_i + f_i = \frac{1}{\gamma_i}(\omega_{0,i}\,{}^2Y_i + {}^2S_i),$$

where ${}^2S_i = \sum_{k=1}^{i} \omega_{k,i}\,{}^2Y_{i-k} + {}^2Y_0/\alpha$.

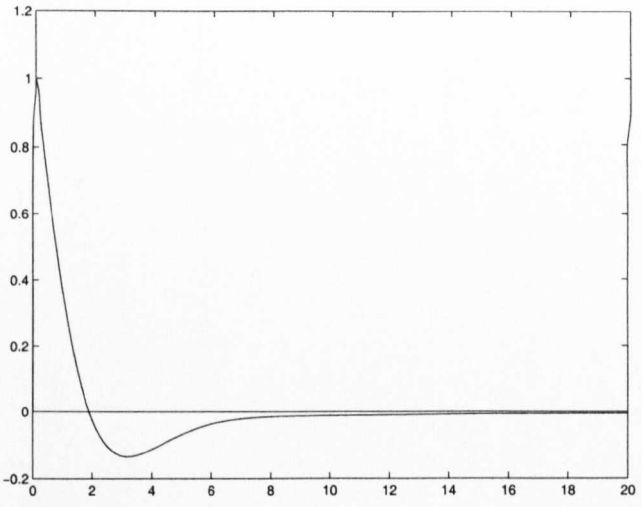

Figure 12.7: Approximate solution to equation (12.10) for $\alpha = 0.3$.

Rearranging gives

$$\begin{pmatrix} 1 & -\frac{h}{2} \\ \gamma_i b & \omega_{0,i} \end{pmatrix} \begin{pmatrix} {}^1Y_i \\ {}^2Y_i \end{pmatrix} = \begin{pmatrix} {}^1Y_{i-1} + \frac{h\,{}^2Y_{i-1}}{2} \\ \gamma_i f_i - {}^2S_i \end{pmatrix}.$$

We give the approximate numerical solution to the example fractional oscillation equation

$$D^{1+\alpha} y = -y,\ y_0 = 1,\ y^{(1)} = 0 \tag{12.10}$$

for $\alpha = 0.3$, 0.5, 0.8 and 0.95 in Figures 12.7, 12.8, 12.9, and 12.10 respectively. Our results are consistent with those given in [8] and are consistent with the known properties of Mittag-Leffler functions [19].

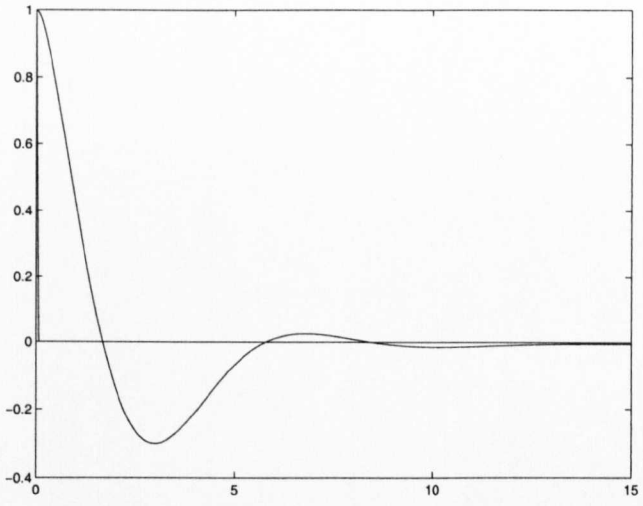

Figure 12.8: Approximate solution to equation (12.10) for $\alpha = 0.5$.

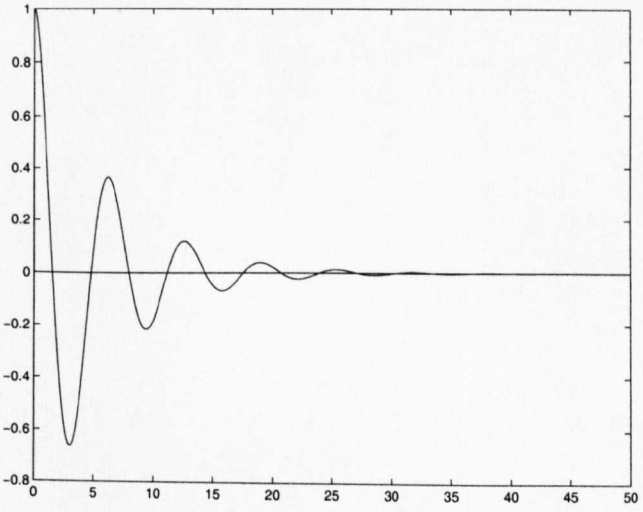

Figure 12.9: Approximate solution to equation (12.10) for $\alpha = 0.8$.

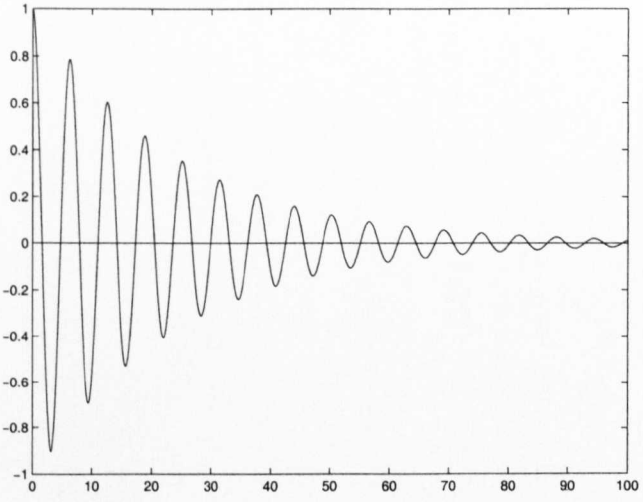

Figure 12.10: Approximate solution to equation (12.10) for $\alpha = 0.95$.

The Basset Equation

As before by the Basset equation we mean any linear fractional differential equation of the form:

$$Dy + b_2 D^\alpha y + b_1 y = f, \tag{12.11}$$

$$y(0) = y_0,\ \alpha \in (0,1),\ b_1 b_0 \neq 0.$$

By the linearity of D and D^α we can rewrite this as

$$\begin{aligned} {}^1Y &= y, \\ {}^2Y &= D^\alpha y, \\ {}^3Y &= Dy. \end{aligned}$$

In matrix notation this gives

$$\begin{pmatrix} D^\alpha & 0 \\ D & 0 \end{pmatrix} \begin{pmatrix} {}^1Y \\ {}^2Y \end{pmatrix} = \begin{pmatrix} {}^2Y \\ f - b_2\, {}^2Y - b_1\, {}^1Y \end{pmatrix}.$$

On discretisation this results in the system

$$\begin{pmatrix} -\omega_{0,i} & \gamma_i \\ \tau_1 & \tau_3 \end{pmatrix} \begin{pmatrix} {}^1Y_i \\ {}^2Y_i \end{pmatrix} = \begin{pmatrix} {}^1S_i \\ F_i + \tau_2\, {}^1Y_{i-1} - \tau_3\, {}^2Y_{i-1} \end{pmatrix},$$

where $F_i = \frac{h}{2}(f_i + f_{i-1})$, $\tau_1 = (1 + \frac{hb_1}{2})$, $\tau_2 = (1 - \frac{hb_1}{2})$, $\tau_3 = (\frac{hb_2}{2})$, and ${}^1S_i = \sum_{k=1}^{i} \omega_{k,i}\, {}^1Y_{i-k} + {}^1Y_0/\alpha$.

To compare our work with that of a previous researcher we use the formulation of the Basset equation given in [44]. Thus we rewrite the Basset equation as

$$Dy + aD^{\alpha}y + y = 1,\ y(0) = 0,\ a = \beta^{\alpha},$$

where $\beta = 9/(1 + 2\chi)$ for some $\chi > 0$.

Our calculated approximate solutions to the Basset equation, for $\chi =$ 0.5, 2, 10, 100 respectively, are given in figures 12.11, 12.12, and 12.13. These appear to agree, up to visual tolerance, with those given in [44] (obtained by an unspecified method).

12.2.5 Numerical Solution of the General Linear Multiterm Equation

For a general linear multiterm fractional equation we consider

$$\sum_{s=0}^{p} c_s D^{\beta_s} y = f, \tag{12.12}$$

where $0 \leq \beta_0 < \cdots < \beta_r < \beta_{r+1} < \cdots < \beta_p$, $c_s \in \mathbb{R}$ and $c_p = 1$. The highest order β_p need not be an integer, however we need to regard every integer order derivative as being present in our equation, if necessary with coefficient 0, so as to use all the initial conditions. We choose to use a

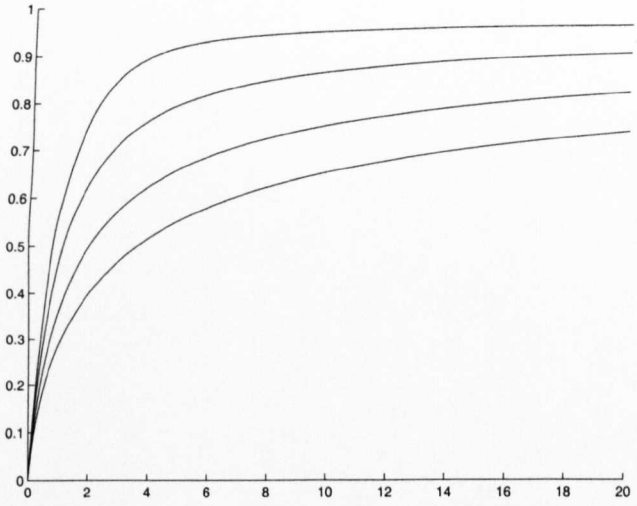

Figure 12.11: Approximate solution to (12.12) for $\alpha = 0.5$.

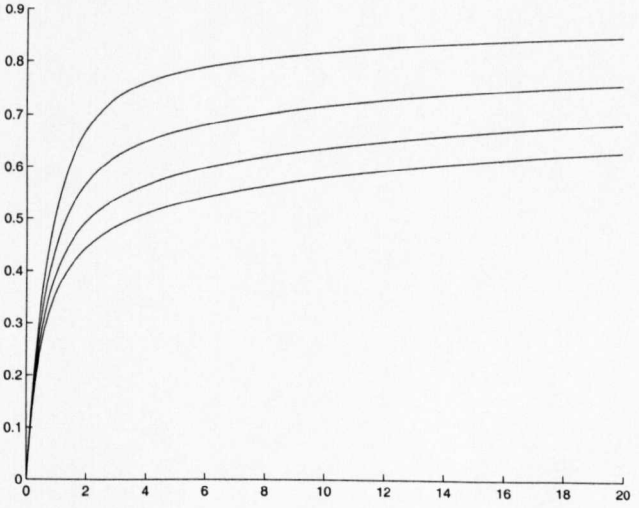

Figure 12.12: Approximate solution to (12.12) for $\alpha = 0.25$.

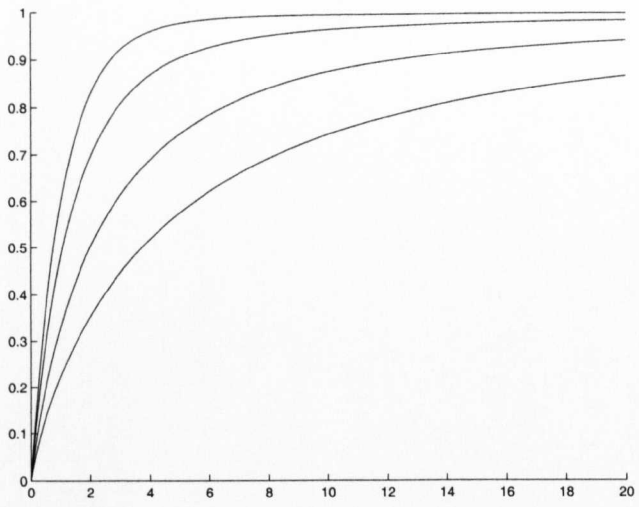

Figure 12.13: Approximate solution to (12.12) for $\alpha = 0.75$.

notation which emphasises this: we collect all the orders within each interval $(j, j+1)$, $j \in \mathbb{Z}$ and so shall use the form

$$b_{0,0}y + \sum_{j=0}^{m} \sum_{r=1}^{n_j} b_{r,j} D^{j+\alpha_{r,j}} y = f, \tag{12.13}$$

where, for $j = 1, 2, \ldots, m-1$, we have $0 < \alpha_{1,j} < \alpha_{2,j} < \cdots < \alpha_{n,j} = 1$ and $0 < \alpha_{1,m} < \alpha_{2,m} < \ldots \alpha_{n_m,m} \leq 1$ and $b_{i,j} \in \mathbb{R}$. Also let $n_0 + \cdots + n_m = p$ and let $P_k = \sum_{j=0}^{k-1} n_j$.

We write equation (12.13) as the system

$$\begin{aligned}
{}^1Y &= y, \\
{}^2Y &= D^{\alpha_1,0}y, \\
&\vdots \\
{}^{1+n_0}Y &= Dy, \\
{}^{2+n_0}Y &= D^{\alpha_1,0}Dy, \\
&\vdots \\
{}^{p+1}Y &= D^{\alpha_{n_m,m}}D^{m-1}y,
\end{aligned}$$

which in matrix form is

$$DY = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & D_m \end{pmatrix} \begin{pmatrix} {}^1Y \\ \vdots \\ {}^{1+n_0}Y \\ \vdots \\ {}^pY \end{pmatrix} = \begin{pmatrix} {}^2Y \\ \vdots \\ {}^{2+n_0}Y \\ \vdots \\ f - \sum_{k=0}^{p-1} c_k \, {}^{1+k}Y \end{pmatrix},$$

where for $k = 1, \dots, m-1$ D_k is a $n_k \times n_k$ matrix of differential operators of the form

$$\begin{pmatrix} D^{\alpha_1,k} & 0 & \ddots & 0 \\ D^{\alpha_2,k} & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ D & 0 & \ddots & 0 \end{pmatrix}.$$

The corresponding expression holds for $k = m$ if $\alpha_{n_m,m} = 1$, which we shall call case 1. If $\alpha_{n_m,m} \neq 1$, which we shall call case 2, we have

$$\begin{pmatrix} D^{\alpha_1,k} & 0 & \ddots & 0 \\ D^{\alpha_2,k} & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ D^{\alpha_{n_m,m}} & 0 & \ddots & 0 \end{pmatrix}.$$

The Numerical Scheme

We discretise the derivatives in D to produce $\bar{D}$. The matrix $\bar{D}$ will consist of square matrix blocks $\bar{D}_k$ along the diagonal which are the discrete analogues of D_k for $k = 1, \ldots, m-1$. The bottom horizontal band of the matrix consists of the matrices $\bar{D}_{m,1}, \ldots, \bar{D}_{m,m}$.

This means we will write the discretised system as

$$\bar{D}Y_i = S_i. \tag{12.14}$$

Where

$$\bar{D} = \begin{pmatrix} \bar{D}_1 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & \bar{D}_2 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 & \bar{D}_2 & 0 \\ \bar{D}_{m,1} & \ldots & \ldots & \ldots & \ldots & \bar{D}_{m,m} \end{pmatrix}, \tag{12.15}$$

and

$$\bar{D}_k = \begin{pmatrix} -^{1,k}\omega_{0,1} & ^{1,k}\gamma_i & 0 & \ldots & \ldots & 0 \\ -^{2,k}\omega_{0,1} & 0 & ^{2,k}\gamma_i & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ -^{n_k-1,k}\omega_{0,1} & 0 & \ddots & 0 & ^{n_k-1,k}\gamma_i & 0 \\ 1 & 0 & \ldots & \ldots & 0 & -\frac{h}{2} \end{pmatrix},$$

and consequently by implication we also have

$$^{k}\bar{Y}_i = \begin{pmatrix} ^{1+p_k}Y_1 \\ \vdots \\ \vdots \\ ^{p_{k+1}}Y_i \end{pmatrix},$$

and

$$ {}^{k}\bar{S}_i = \begin{pmatrix} {}^{1+p_k}S_1 \\ \vdots \\ \vdots \\ {}^{p_{k+1}}S_i \end{pmatrix}, $$

where

$$ {}^{j+p_k}S_i = \sum_{r=0}^{i-1} {}^{j+p_k}\omega_{i-r,i}\, {}^{1+p_k}Y_r + \frac{{}^{1+p_k}Y_0}{{}^{j+p_k}\beta - (k-1)}, $$

for $j = 1, \ldots, n_k - 1$ and

$$ {}^{p_{k+1}}S_i = {}^{1+p_k}Y_{i-1} + \frac{h}{2}\, {}^{p_{k+1}}Y_{i-1}. $$

Discretising for $k = m$ and collecting terms together gives the first $n_m - 1$ rows of $\bar{D}_{m,m}$ as

$$ \begin{pmatrix} -{}^{1,m}\omega_{0,i} & {}^{1,m}\gamma_i & 0 & \cdots & 0 \\ -{}^{2,m}\omega_{0,i} & 0 & {}^{2,m}\gamma_i & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ -{}^{n_m-1,m}\omega_{0,i} & 0 & \ddots & {}^{n_m-1,m}\gamma_i & 0 \end{pmatrix}. $$

For the last row of $\bar{D}$ we have two cases to consider depending on whether $D^{\alpha_{n_m,m}} \in (0,1)$ or $D^{\alpha_{n_m,m}} = 1$.

The last row of D is determined by the equation

$$ D^{\alpha_{n_m,m}}\, {}^{1+p_m}Y_i = f - \sum_{s=1}^{p} c_{s-1}\, {}^{s}Y_i. \tag{12.16} $$

When $\alpha_{n_m,m} \neq 1$ on discretisation (12.16) becomes

$$ \frac{1}{{}^{n_m,m}\gamma_i}\Big(\sum_{k=0}^{i} {}^{n_m,m}\omega_{i,k}\, {}^{1+p_m}Y_{i-k} + \frac{{}^{1+p_m}Y_0}{\alpha^{n_m,m}}\Big) = f_i - \sum_{s=1}^{p} c_{s-1}\, {}^{s}Y_i. $$

Rearranging this

$$-{}^{n_m,m}\omega_{i,0}\,{}^{1+p_m}Y_i + {}^{n_m,m}\gamma_i \sum_{s=1}^{p} {}^{s}Y_i = {}^{n_m,m}S_i,$$

where

$${}^{n_m,m}S_i = {}^{n_m,m}\gamma_i f_i - \sum_{k=1}^{i} {}^{n_m,m}\omega_{i,k}\,{}^{1+p_m}Y_{i-k} - \frac{{}^{1+p_m}Y_0}{\alpha^{n_m,m}}.$$

Then for $k = 1, \ldots, m-1$, $\bar{D}_{m,k}$ has all zero entries for the first $n_m - 1$ rows. The last row of $\bar{D}_{m,1}$ is given by

$$\left({}^{n_m,m}\gamma_i c_0 \quad \ldots \quad {}^{n_m,m}\gamma_i c_{n_0}\right).$$

For $k = 2, \ldots, m-1$ the last row of $\bar{D}_{m,k}$ is given by

$$\left({}^{n_m,m}\gamma_i c_{p_k+1} \quad \ldots \quad {}^{n_m,m}\gamma_i c_{p_k+1}\right).$$

For $\bar{D}_{m,m}$ the last row is given by

$$\left({}^{n_m,m}\gamma_i c_{p_m+1} - {}^{n_m,m}\omega_{i,0} \quad \ldots \quad {}^{n_m,m}\gamma_i c_{p_k+1}\right).$$

When $\alpha_{n_m,m} = 1$ on discretisation (12.16) becomes

$${}^{1+p_m}Y_i = {}^{1+p_m}Y_{i-1} + \frac{h}{2}(f_i + f_{i-1} - \sum_{s=1}^{p} c_{s-1}\,({}^{s}Y_i + {}^{s}Y_{i-1})).$$

Rearranging this gives

$${}^{1+p_m}Y_i + \frac{h}{2}\sum_{s=1}^{p} c_{s-1}\,{}^{s}Y_i = {}^{n_m,m}S_i,$$

where

$${}^{n_m,m}S_i = {}^{1+p_m}Y_{i-1} + F_i - \frac{h}{2}\sum_{s=1}^{p} c_{s-1}\,{}^{s}Y_{i-1}.$$

As before $\bar{D}_{m,k}$, for $k = 1, \ldots, m-1$, has all zero entries for the first $n_m - 1$ rows. For $\bar{D}_{m,1}$ the last row is given by

$$\left(\tfrac{h}{2}c_0 \quad \cdots \quad \tfrac{h}{2}c_{n_0}\right),$$

and for $k = 2, \ldots, m-1$ the last row of $\bar{D}_{m,k}$ is given by

$$\left(\tfrac{h}{2}c_{1+p_k} \quad \cdots \quad \tfrac{h}{2}c_{p_{k+1}}\right),$$

and the last row of $\bar{D}_{m,m}$ is given by

$$\left(1 + \tfrac{h}{2}c_{1+p_m} \quad \tfrac{h}{2}c_{2+p_m} \quad \cdots \quad \tfrac{h}{2}c_p\right).$$

12.2.6 Conclusion

In this section we have shown that linear multiterm fractional differential equations are as amenable to being solved as systems of equations as are linear ordinary differential equations. The order of convergence of the underlying numerical methods are conserved and the resulting algorithms are easily programmable.

Chapter 13

Conclusions

13.0.7 Power-law Decay and Chaos

We have exploited the power-law properties of the kernel of the fractional differential operator and the existence of an invariant scale to develop our quicker algorithms. Fractals are another area where power-law properties play a significant role and fractals are important in non-linear dynamics. A question that this prompts is how would the long term behaviour of the models, sets of coupled differential equations, of chaos alter if one, or some, of the differential operators in the set of differential equations were to be replaced with fractional differential operators

In a linear system a perturbation introduces a permanent set whereas non-linear systems are dissipative. Fractional differential operators have something of a 'fading out' property'. For instance if we move the pepper pot one square away from the salt pot, on a chequered table cloth, we expect it to stay there, in linear space time, whereas in fractional space time we would expect the pepperpot to 'shrink back' to its original position.

13.0.8 Collocation

We have used the fractional generalisations of the trapezium rule and linear multistep methods. A fractional version of semi-implicit Runge-Kutta methods, or collocation, also exists. It was remarked, by the attending co-author of [33], that collocation methods are regarded as being generally the most accurate methods for approximating the solutions of second kind Volterra integral equations, although a satisfactory convergence has not yet been given. The method given in [8] is a specialisation of collocation to fractional differential equations.

We have seen that our ideas work well with the fractional trapezium rule. When applying these ideas, in the nested mesh variant, to the fractional linear multistep methods we obtained the right sort of convergence behaviour. However a more violent reaction to the step length change occurred than was the case for the fractional trapezium rule. We also had a problem of interpreting the starting weights for the 'off grid' iteration. In addition we have the colossal cost of computing the weights for all but the simplest methods.

Collocation methods should allow us to obtain the same orders of convergence as the FLMM but with other parts of the algorithm of the same order as the fractional trapezium rule. That is we should be able to calculate the weights for the fractal sum implementation with the same order of accuracy. Additionally we may attempt to use collocation to approximate the kernel of a linear multiterm fractional differential equation.

Reference is frequently made to the use of FFTs in calculating the weights for FLMM. The usefulness of this depends on how much difference there is in

the cost of doing multiplication's compared to doing additions. Using FFTs, for two vectors of length n, we have $2n-2$ additions and 1 multiplication where conventionally we would have n multiplications and $n-1$ additions, which in both cases is $2n-1$ floating point operations.

13.0.9 Special Functions

When considering the numerical approximation of the solution of the non-linear fractional differential equation of the form

$$D^{\alpha}y = \lambda y^2,\ \alpha > 0,\ \lambda \in \mathbb{C}, \tag{13.1}$$

we found convergent behaviour for $\alpha \in (0,1)$ and divergent behaviour for $\alpha \in (1,2)$.

For $\alpha = 1$ we can calculate the exact solution to equation 13.1 and find that we will have one pole in the complex plane. For $\alpha = 2$ the differential equation is then of the same form as the differential equation satisfied by the Weierstrass elliptic function $\mathcal{P}$ and therefore the solution to (13.1) is asymptotic to $\mathcal{P}$ [31] and the divergent behaviour observed in our attempts at approximating (13.1) is to be expected.

This leads us to ask is there an interesting class of special functions, which solves (13.1), waiting to be discovered.

13.0.10 Time Dilation

If we consider a planet P_c performing a circular orbit, of period T about a point S, then its orbital velocity will be constant. If we consider a planet P_e in an elliptical orbit, again of period T about a point S, we know by one of Kepler's laws (equal areas in equal times) that its velocity must vary.

Therefore by special relativity the rate at which time passes, for one orbit, on the two planets, relative to passage of time at the point S, will be different. We therefore might expect asymmetry to affect how time propagates in a system.

If we consider a long chain polymer we may regard it is an articulation of a very large number of thin cylindrical elements of radius r and length l where $l >> r$. If a lump of material composed of long chain polymers has been undisturbed, for a long period, we might expect the orientation of these elements to be randomly distributed, that is that asymmetry is symmetrically distributed. If we now deform it we would expect the distribution of the elements to be in the direction of the deformation. Then the distribution of asymmetry is no longer symmetrically distributed.

If we speculate that each element can be regarded as an infinitesimal 'time' element, like a magnetic element, summing over a symmetrical distribution would cause the asymmetries to cancel out, however if they had a tendency to line up in the same direction a slowing down of time might occur.

13.0.11 Non-standard Analysis

All the developments of the fractional calculus referred to in this thesis give a development in terms of ϵ-δ methods. A development based on a non-standard model of arithmetic may lead to some useful insights.

For instance the original motivation for the nested sum algorithm comes from using a variation in step length size to model the behaviour of a T shaped plate using partial differential equations. A smaller step length being

used to model the diffusion of heat or the propagation of cracks in the 'arm pits' of the T.

Let $\mathbb{R}^*$ be an extension of the real numbers containing infinitely large and infinitely small elements. Let $\delta > 0 \in \mathbb{R}^*$ be infinitely small. Then the D^δ derivative will have a kernel of the form $(t-s)^\delta$ which will retain its scaling properties and will be infinitely close to 1 for all $t \in \mathbb{R}^+$. That is in $\mathbb{R}^*$ the kernel will tend to zero whilst in $\mathbb{R}$ its standard part will always be unity. Thus using non-standard arithmetic we can model the features of the T which inspired our results by the standard part of the numbers and the fractional scaling property, which justifies our approach, by appealing to features of the extended model of arithmetic.

13.0.12 Reaction of Bone to Violent Shock

A discovery of modern material science is that larger objects (composed of the same material and of the same shape) are more brittle than smaller objects due to the propagation of micro-cracks. Thus *the bigger they are the harder they fall* is a faster than linear relationship.

Violence between individuals is a feature of human society. There are many results in mathematical physics to the effect that maximum power transfer occurs under symmetrical conditions. We might therefore speculate that violence between symmetrical individuals is more destructive and that the bigger the individuals are the more physically damaging that violence is.

One of the conclusions of material science is that for many applications a hollow tube is just as strong as a solid tube and therefore using a hollow tube saves on material. There is no point in telling this to the birds since

they already know it. This suggests that the equivalent bone in a small scale person could suffer a larger deflection (in terms of curvature) than for a large person for two reasons: firstly the slower onset of tension in the opposite side of the bone and secondly less intrinsic bone brittleness.

Appendix A

The Tautochrone

In the problem of the tautochrone we ask what is the shape of a curve C such that the time taken for a frictionless bead to slide down it is independent of the beads starting point, see figure A.1.

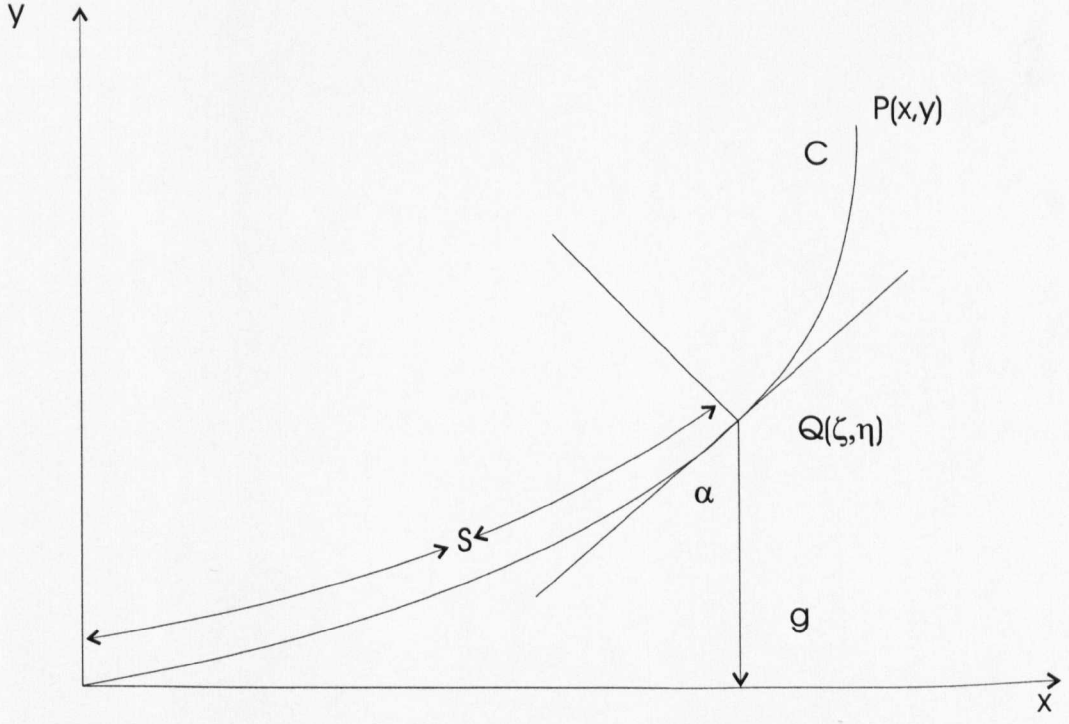

Figure A.1: The tautochrone problem

Solution of the Tautochrone

At the point Q the acceleration of the bead along the curve is

$$\frac{d^2s}{dt^2} = -g\cos\alpha, \tag{A.1}$$

and

$$\cos\alpha = \frac{d\eta}{ds}, \tag{A.2}$$

which gives

$$\frac{d^2s}{dt^2} = -g\frac{d\eta}{ds}. \tag{A.3}$$

We now multiply by $\frac{ds}{dt}$ and integrate by parts to obtain

$$(\frac{ds}{dt})^2 = -2g\eta + k. \tag{A.4}$$

Now since the bead starts from rest $\frac{ds}{dt} = 0$, at $\eta = y$ we must have $k = 2gy$ and, taking the negative square root since s reduces as t increases, we have

$$\frac{ds}{dt} = -\sqrt{2g(y-\eta)}. \tag{A.5}$$

Thus the time of descent T from P to O is given by

$$T = -\frac{1}{\sqrt{2g}}\int_P^O \frac{1}{\sqrt{y-\eta}}ds. \tag{A.6}$$

The arc length s is a function of η so $s = h(\eta)$ for some function h. Assuming h is differentiable we can write the previous equation as

$$T = -\frac{1}{\sqrt{2g}}\int_y^O \frac{h'(\eta)}{\sqrt{y-\eta}}d\eta. \tag{A.7}$$

Writing $h'(y)$ as $f(y)$ and generalising the differential operator notation to the fractional case we can write this as

$$\frac{\sqrt{2g}}{\Gamma(\frac{1}{2})}T = D^{-\frac{1}{2}}f(y). \tag{A.8}$$

If we assume that f has bounded integrals and derivatives of all orders then we can rewrite this as

$$D^{\frac{1}{2}}\sqrt{\frac{2g}{\pi}}T = f(y), \tag{A.9}$$

and since

$$D^{\frac{1}{2}}T = \frac{T}{\sqrt{\pi y}}, \tag{A.10}$$

we have

$$f(y) = \frac{\sqrt{2g}}{\pi}Ty^{-\frac{1}{2}}. \tag{A.11}$$

To find the equation of the curve C we have

$$f(y) = h'(y) = \frac{ds}{dy} = (1 + (\frac{dx}{dy})^2)^{\frac{1}{2}}, \tag{A.12}$$

so

$$\frac{dx}{dy} = \sqrt{f^2 - 1}, \tag{A.13}$$

therefore

$$x = \int_0^y \sqrt{\frac{2gT^2}{\pi^2\eta} - 1}d\eta + c. \tag{A.14}$$

Setting $x = 0$ at $y = 0$ gives $c = 0$, let $a = gT^2/\pi^2$ then

$$x = \int_0^y (\frac{2a}{\eta} - 1)^{\frac{1}{2}}d\eta. \tag{A.15}$$

Changing the variables of integration to $\eta = 2a\sin^2\theta$ and $d\eta/d\theta = 4a\sin\theta\cos\theta$ gives

$$\begin{aligned} x &= \int_0^{\sin^{-1}\sqrt{\frac{y}{2a}}} (\frac{\cos^2\theta}{sin^2\theta})4a\sin\theta\cos\theta d\theta, \\ &= \int_0^{\sin^{-1}\sqrt{\frac{y}{2a}}} \cos^2\theta d\theta, \\ &= 4a\left[\frac{\theta}{2} + \frac{1}{4}\sin 2\theta\right]_0^{\sin^{-1}\sqrt{\frac{y}{2a}}}. \end{aligned} \tag{A.16}$$

Writing $\beta = \sin^{-1}\sqrt{y/2a}$ we have

$$x = 2a(\beta + \frac{1}{2}\sin 2\beta), \tag{A.17}$$

and

$$y = 2a\sin^2\beta. \tag{A.18}$$

If we now let $2\beta = \theta$, and recall that $\cos\theta = \cos^2\theta - \sin^2\theta = 1 - 2\sin^2\theta/2$, gives the parametric representation of C as

$$\begin{aligned} x &= a(\theta + \sin\theta), \\ y &= a(1 - \cos\theta), \end{aligned}$$

which is the parametric form of the equation for a cycloid.

Appendix B

Classical Viscoelastic Theory

B.1 Motivation

We give a very brief and basic survey of viscoelastic theory. The interpretation of viscoelasticity given here is a synthesis of [47] and [39]. For a review of the basic physics of materials see Appendix C.

B.2 Classical viscoelastic Theory

B.2.1 Basic Concepts

In classical linearised elasticity

$$stress\ (in\ a\ sheared\ body) \propto amount\ of\ shear \tag{B.1}$$

The Navier-Stokes theory of viscosity asserts

$$shearing\ stress \propto rate\ of\ shear \tag{B.2}$$

If, in a material, these effects are not further complicated by behaviour that is unlike either elasticity or viscosity the material is called viscoelastic. We will denote stress by $\sigma(t)$ and strain by $\kappa(t)$.

B.2.2 Stress Relaxation

We consider the behaviour of a slab of material in simple shearing motion, see figure B.1 . The slab is to be regarded as so thin that inertial effects can be ignored.

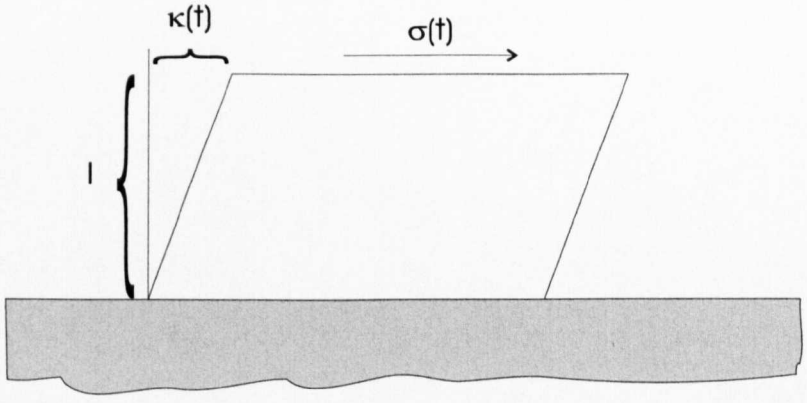

Figure B.1: Slab of material subjected to a shearing stress.

Then the slab can be regarded as homogeneously deformed, with the amount of shear strain $\kappa(T)$ variable in time. Let $\sigma(t)$ be the shearing stress per unit area on the slab. In the case of an elastic material the stress history has the form

$$\sigma(t) = \sigma_0 H(t) \tag{B.3}$$

where $H(t)$ is the Heaviside step function. If the material were an ideal viscous fluid the stress would be instantaneously infinite during the step, and then zero for all time afterwards, like a Dirac delta function, $\delta(t) = H'(t)$, figure B.2.

However the stress usually decreases from its initial value quite rapidly at first, and later more gradually, approaching some limiting value $\sigma(\infty)$, figure

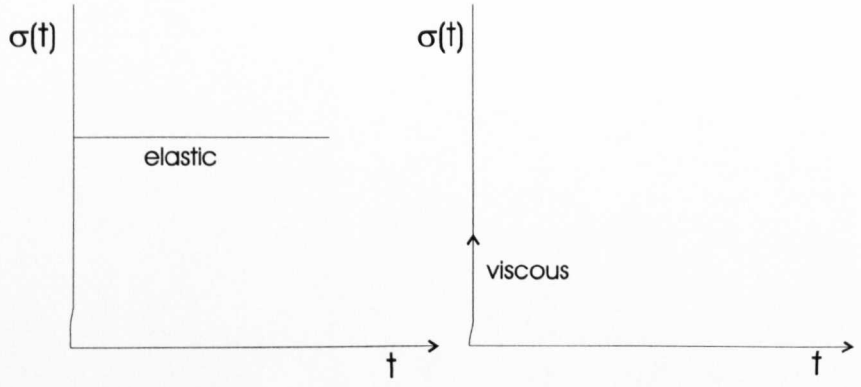

Figure B.2: Ideal elastic and viscous behaviours.

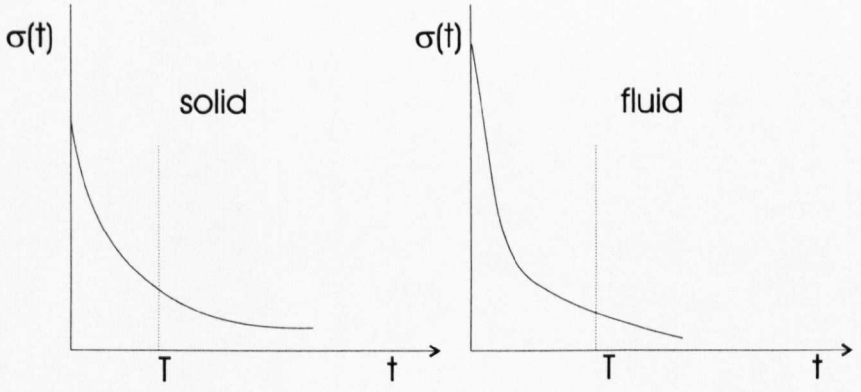

Figure B.3: Behaviour of solids and fluids.

B.3. If the limiting value is not zero we are likely to call the material a solid. If the limiting value is zero, and the approach to zero sufficiently rapid, we call the material a fluid. Let T be the relaxation time. We call materials viscoelastic, and use appropriate mathematical models, when the relaxation time and period of observation are not astronomically different.

B.2.3 Creep

Now suppose that a slab is subjected to a one step stress history $\sigma(t) = \sigma_0 H(t)$. The response of an elastic solid would be $\kappa(t) = \kappa_0 H(t)$, constant shear for $t > 0$. In a viscous fluid, the sheer would increase at a constant rate $\chi(t) = \frac{\sigma_0 t}{\eta}$, the coefficient η being the viscosity. Viscoelastic theory recognises

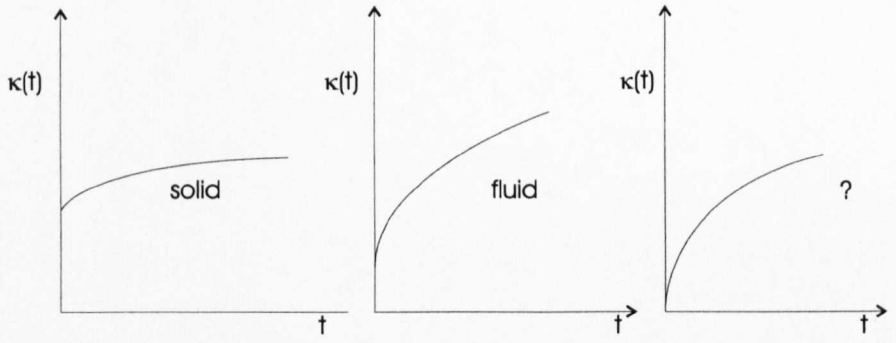

Figure B.4: Three models of shear response.

more refined observations which show departures from these idealisations. The shear at first jumps, so far as anyone can tell, so that the instantaneous response is elastic. The shear then continues to increase, but at lower and lower rates, figure B.4.

B.2.4 Response Functions

Let $R(\kappa, t)$ be the stress relaxation function, the stress t units of time after application of a shear step of size κ. Let $C(\sigma, t)$ be the creep function, the shear t units of time after application of a stress σ. R and C are zero for $t < 0$. If the material is isotropic then by consideration of symmetry R is an odd function of κ and C. Hence assuming smooth dependence and supposing

that κ and σ are small we have

$$R(\kappa, t) = G(t)\kappa + O(\kappa^3), \tag{B.4}$$

and

$$C(\sigma, t) = J(t)\sigma + O(\sigma^3). \tag{B.5}$$

The values of these functions at $t = 0+$ are denoted G_g and J_g (g for glass), and the values at $t = \infty$ are G_e and J_e (e for equilibrium or elastic depending on the author), provided that these values exist. If $J(t)$ tends to increase like $\frac{t+T}{\eta_0}$ for large t, η_0 is called the steady state shearing viscosity and T is the mean relaxation time, figure B.5.

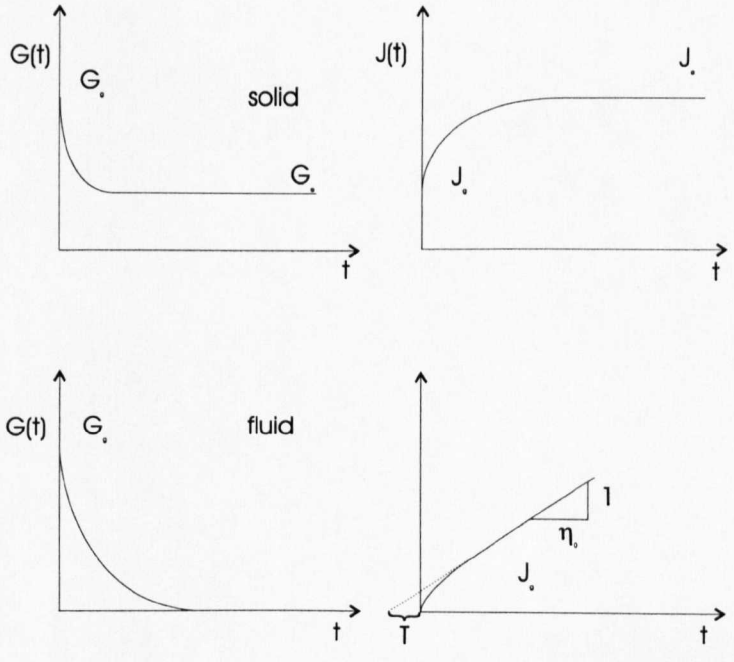

Figure B.5: Response behaviour for solids and fluids.

Immediately after application of a step strain or stress, the response is independent of whether it is the strain or stress which is to be held constant in

the future. Hence at time 0+, $\sigma = G_g\kappa$ and $\kappa = J_g\sigma$. Thus we find that the initial values of G and J are reciprocal:

$$J_g G_g = 1. \tag{B.6}$$

If the stress and strain approach limiting values after a long time for viscoelastic substances it is irrelevant which one was held absolutely constant for all +ve time. Thus in the limit as $t \to \infty$ both $\sigma = G_e\kappa$ and $\kappa = J_e\sigma$ are valid. Hence

$$J_e G_e = 1. \tag{B.7}$$

Apparently $J(t)$ and $G(t)$ are roughly reciprocal at all times, although exactly so only in the limiting cases. The reciprocal relations for the limits can be viewed as the assumptions that the instantaneous response and equilibrium response are elastic.

B.2.5 Models

A spring is an ideal elastic element obeying the linear force-extension, Hooke's law, $\sigma = G\kappa$, figure B.6. Its relaxation modulus is $G(t) = GH(t)$, and its creep compliance is $J(t) = JH(t)$. Here $J = \frac{1}{G}$.

A dashpot is an ideal viscous element that extends at a rate proportional to the applied force, $\dot{\kappa} = \frac{\sigma}{\eta}$, figure B.7. Hence $J(t) = \frac{tH(t)}{\eta}$ and $G(t) = \eta\delta(t)$.

Springs and dashpots can be combined in different ways to produce alternative models of viscoelastic behaviour, figure B.8.

When two elements are combined in series, their compliances are additive. Thus the Maxwell model consists of a spring and dashpot in series and has

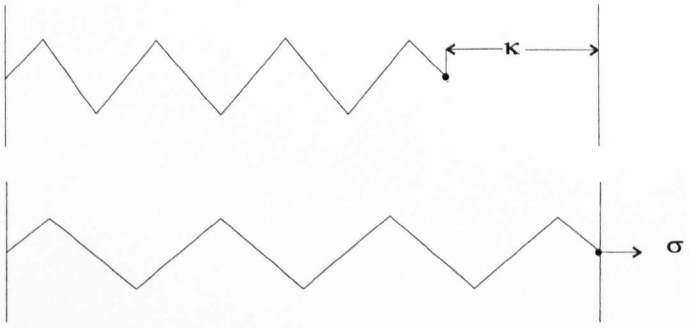

Figure B.6: Spring or Hooke's model of an ideal elastic element.

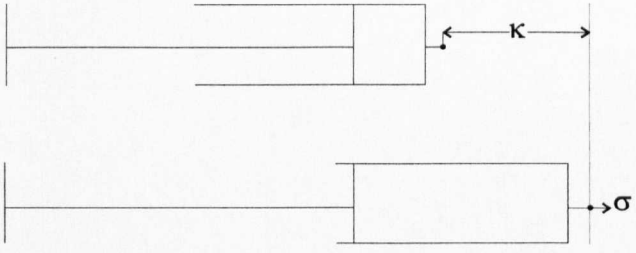

Figure B.7: Dashpot or Newton's model of an ideal viscous element.

creep compliance

$$J(t) = (J_g + \frac{t}{\eta})H(t). \tag{B.8}$$

When two elements are combined in parallel, their moduli are additive. The Kelvin-Voigt model consisting of a spring and dashpot in parallel has the modulus

$$G(t) = G_e H(t) + \eta\delta(t). \tag{B.9}$$

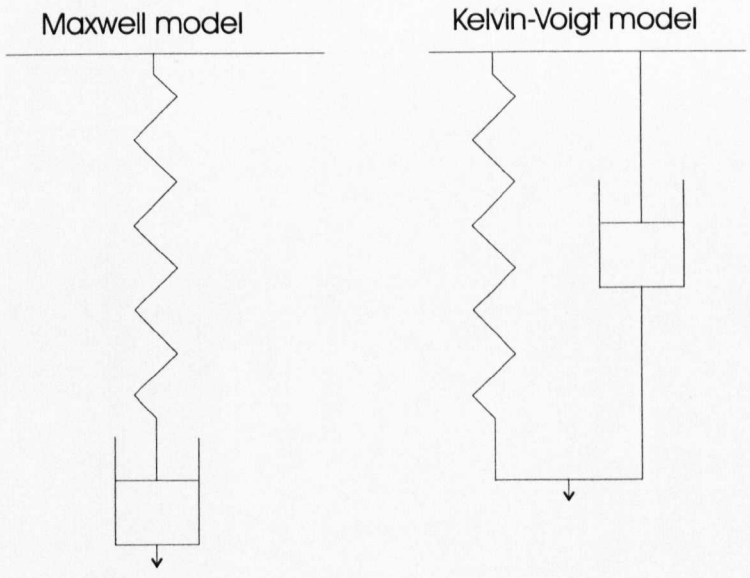

Figure B.8: Simple dashpot and spring models of viscoelastic behaviour.

Appendix C

Properties of Materials

C.0.6 Basic Concepts

Strength how great an applied force a material can withstand before breaking.

Stiffness the opposition of a material to being distorted by having its size or shape changed.

Ductility (or workability) relates to the ability of the material to be hammered, pressed, bent, rolled, cut or stretched into useful shapes.

Toughness does not crack readily, i.e. is not brittle.

At normal temperatures steel has all four of these basic properties whilst custard has none. When acted on by a force the deformation produced in a sample of material depends on

1. the nature of the material
2. the stretching force
3. the cross sectional area of the sample

4. the original shape of the sample

Two important concepts necessary for understanding viscoelasticity are stress and strain.

Stress is the force acting on the unit cross-section area and for a force F and area A it equals $\frac{F}{A}$. The units of stress are Newtons per square metre (Nm^{-1})

Strain is the extension of unit length if e = extension and l = original length then strain $=\frac{e}{l}$. Strain is a ratio and has no units.

A stress which causes an increase of length puts the sample in tension, thus we talk about tensile stress and tensile strain. The stress strain graph depends not only on the material but also on previous uses and methods of manufacture.

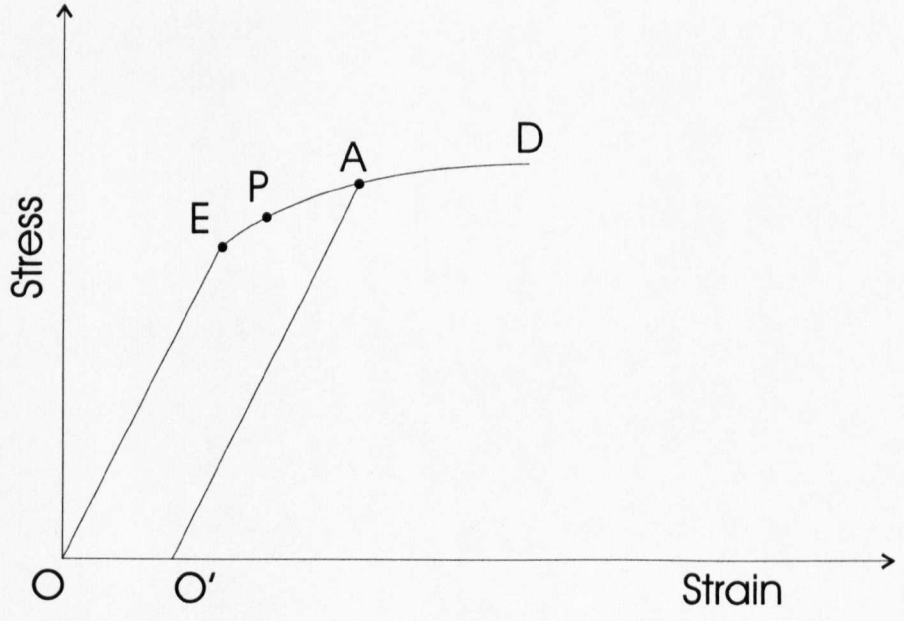

Figure C.1: Plot of stress against strain.

C.0.7 Elastic Deformation

The first part of the graph, given in figure C.1 from OE is a straight line through the origin and here strain is directly proportional to stress. Over this range the material suffers elastic deformation, i.e. it returns to its original length when the stress is removed and therefore none of the original extension remains.

C.0.8 Plastic Deformation

As the stress increases the graph become non-linear but the deformation remains elastic until at a certain stress corresponding to point P, called the yield point, permanent or plastic deformation starts. The material retains some of its extension if the stress is removed. On reducing the stress at A the specimen recovers along AO′ where AO′ is almost parallel to OE. OO′ is the permanent plastic extension produced. If the stress is reapplied, the curve O′AD is followed, At D the specimen develops one or more 'waists' and ductile fracture occurs at one of them. The stress at D is the greatest that the material can bear and is called the breaking stress or ultimate tensile stress. The specimen appears to 'give' at P, and over the plastic region a given stress increase produces a greater increase of strain than previously. Nonetheless it still opposes deformation and any increase of strain requires increased stress. Beyond P the material is said to work harden or strain harden.

Material	Young's mod. $(\times 10^{10} Nm^{-2})$
steel	21
copper	13
glass	7
polythene	0.5
rubber	0.005

Table C.1: Approximate values of E

C.0.9 Young's Modulus

The stress-strain curve for the stretching of metals and some other materials such as glass is linear over almost all the elastic region. This statement is known as Hooke's law.

$$\frac{tensile\ stress}{tensile\ strain} = const. \tag{C.1}$$

This constant, denoted by E, is called Young's modulus. Its value depends on the nature of the material and not on the dimensions of the sample. Materials with large E resist elastic deformations strongly and a large stress is required to produce a small strain. If a stretching force F acting on a wire of cross-sectional area A and original length l causes an extension e we can write

$$E = \frac{tensile\ stress}{tensile\ strain} = \frac{\frac{F}{A}}{\frac{e}{l}} = \frac{Fl}{Ae} \tag{C.2}$$

C.0.10 Fatigue

This may cause fractures, often with little or no warning, and happens when a metal is subjected to a large number of cycles of varying stress even if the maximum value of the stress could be applied steadily with complete safety.

For many ferrous metals there is a safe stress variation below which failure will not occur even for an infinite number of cycles. With other materials 'limited life' design only is possible.

C.0.11 Creep

In general, for metals, this occurs at high temperature and results in the metal continuing to deform as time passes, even under constant stress. Some low melting point metals can creep at low temperatures, e.g. lead sheeting on church roofs has to be replaced periodically. Creep is a particular feature of viscoelastic substances.

C.0.12 Elastic Moduli

All deformations of a body whether stretches, compressions, bends or twists can be regarded as consisting of one or more of three basic types of strain. For many materials experiment shows that provided the elastic limit is not exceeded

$$\frac{tensile\ stress}{tensile\ strain} = const. \tag{C.3}$$

This is a more general statement of Hooke's law. The constant is called an elastic modulus of the material for the type of strain under consideration. There are three moduli, one for each kind of strain.

1. Young's Modulus (E) as previously explained.

2. Rigidity Modulus (G) in this case the strain involves a change of shape without a change of volume.Thus if a tangential force F is applied along the top surface, of area A, of a rectangular block of materials fixed to

the bench, the block suffers a change of shape and is deformed so that the front and rear faces become parallelograms. The shear stress is $\frac{F}{A}$ and angle α taken as a measure of the strain produced. The rigidity modulus is defined as

$$G = \frac{shear\ stress}{shear\ strain} = \frac{F}{\alpha x}.$$

When wire is twisted , a small square on the surface becomes a rhombus, and an example of a shear strain. G can be found from experiments on the twisting of wire. If a spiral spring is stretched, the wire itself is not extended but it is twisted, i.e. sheared. The extension thus depends on the rigidity modulus of the material as well as on the dimension of the spring.

3. Bulk Modulus (K) if a body of volume V, as in Figure C.2, is subject to an increase of external pressure δp which changes its volume by δV, the deformation is a change of volume without a change of shape. The bulk modulus stress is δp, i.e. an increase in $\frac{force}{unit\ area}$ and the bulk strain $\frac{\delta V}{V}$, i.e. a change of $\frac{volume}{original\ volume}$. The bulk modulus K is defined by

$$K = \frac{bulk\ stress}{bulk\ strain} = \frac{-\delta p}{\frac{\delta V}{V}}. \qquad \text{(C.4)}$$

The - sign is introduced to make K +ve since δV is -ve. Solids have all three moduli, liquids and gases only have K. All moduli have the same units Nm^{-2}.

C.0.13 Strain

This has two classifications dilation changes the volume but not the shape
deviatoric changes the shape but not the volume

External forces can be applied to a body in two ways: either, as with gravity

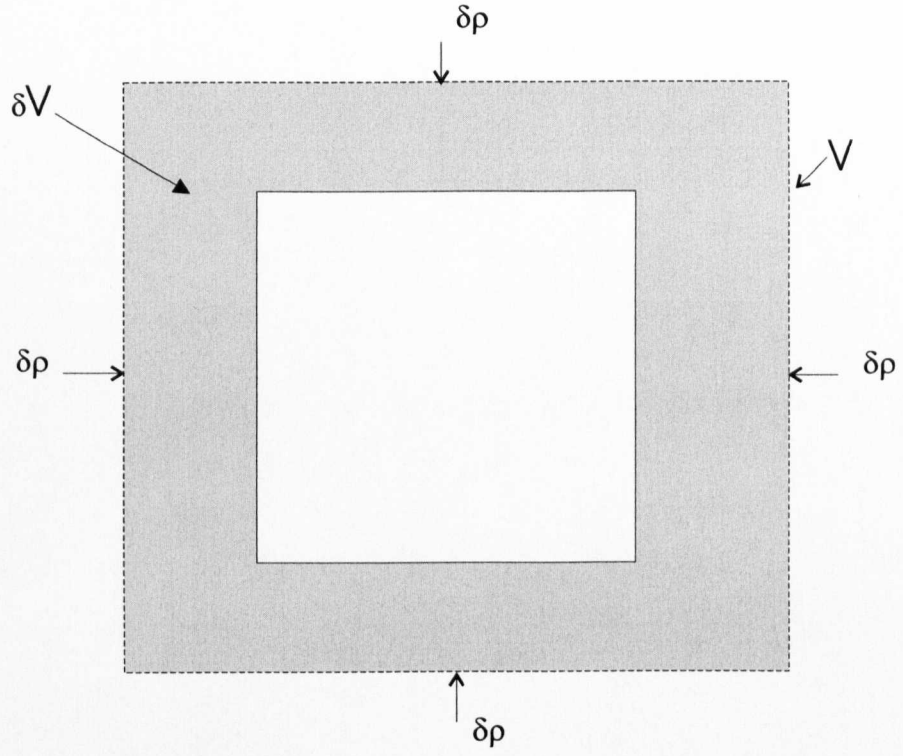

Figure C.2: Cross section of a body subject to an increase in external pressure.

and inertia, as body forces which act directly on the particles of the body; or as surface and contact forces which act directly on the particles at the surface but only indirectly, through forces transmitted along the network of bonds, on those inside the body.

Bibliography

[1] Abel N., *Solution de Quelques Problemes a l'aide d'Integrals Definites*, Oeuvers Completes, Volume 1, 1881.

[2] Humbert P. and Agarwal R.P., *Sur la fonction de Mittag-Leffler et quelquesunes de ses generalisations*, Bulletin des Sciences Mathematiques, vol. 77, no. 10, 180-185, 1953.

[3] Apostol T.M., *Mathematical Analysis* Addison-Wesley, 1977.

[4] Bagley R.L. and Torvik P.J., *On the appearance of the fractional derivative in the behaviour of real materials*, J. Appl. Mech., 51, 294-298, 1984

[5] Baker C.T.H. and Derakhshan M.S., *FFT Techniques in the Numerical Solution of Convolution Equations*, Journal of Computational and Applied Mathematics, 20, 1987.

[6] Baker C.T.H. and Derakhshan M.S., *The Use of NAG FFT Routines in the Construction of Power Series Used in Fractional Quadrature Rules*, Numerical Analysis Report No. 115, Manchester Centre for Computational Mathematics.

[7] Baker C.T.H. and Derakhshan M.S., *A Code for Fast Generation of Quadrature Rules with some Special Properties*, Numerical Analysis Report No. 121, Manchester Centre for Computational Mathematics.

[8] Blank L., *Numerical treatment of differential equations of fractional order*, Numerical Analysis Report No. 287, Manchester Centre for Computational Mathematics, 1996.

[9] Bourbaki N., *Elements of the History of Mathematics*, Springer-Verlag, 1994.

[10] H. Brunner and P.J. van der Houwen, *The numerical solution of Volterra equations*, North-Holland, 323-407, 1986.

[11] Caputo M., *Linear Models of Dissipation whose Q is almost Frequency Independent - II*, Geophysical Journal Royal Astronomical Society, 13, 1967.

[12] Diethelm K. *An Algorithm for the Numerical Solution of Differential Equations of Fractional Order*, Electronic Transactions on Numerical Analysis, Kent State University, 5, 1-6, 1997.

[13] Diethelm K., *Generalised Compound Quadrature Formulae for Finite-part Integrals*, IMA Journal of Numerical Analysis 17, 479-493, 1997.

[14] Diethelm K. and Freed A.D., *On the solution of non-linear fractional order differential equations used in the modelling of viscoplasticity*, Scientific Computing in Chemical Engineering II: Computational Fluid

Dynamics, Reaction Engineering, and Molecular Properties, Springer-Verlag, Heidelberg, 217-224, 1999.

[15] Diethelm K., Ford N.J. and Freed A.D., *A Predictor-Corrector Approach for the Numerical Solution of Fractional Differential Equations*, to appear. 2001.

[16] Diethelm K. and Luchko Y., *Numerical Solution of Linear Multi-Term Differential Equations of Fractional Order*, J. Comput. Anal. Appl., to appear.

[17] Diethelm K. and Walz G., *Numerical Solution of Fractional Order Differential Equations by Extrapolation*, Numerical Algorithms 16, 231-253, 1997.

[18] Elliot D., *An Asymptotic Analysis of Two Algorithm for Certain Hadamard Finite-part Integrals*, IMA J. Numerical Analysis 13, 445-462, 1993.

[19] Erdelyi A., *Higher Transcendental Functions Volume 3*, McGraw-Hill, 1954.

[20] Ford N.J. and Diethelm K., *Analysis of Fractional Differential Equations*, J. Math. Anal. Appl., to appear.

[21] Ford N.J. and Diethelm K., *Numerical solution of the Bagley-Torvik equation*, BIT, 101-116, 2000.

[22] Ford N.J. and Diethelm K., *The numerical solution of linear and non-linear fractional differential equations involving fractional derivatives of several orders* , to appear.

[23] Ford N.J. and Simpson A.C., The numerical solution of linear multi-term fractional differential equations: systems of equations, submitted to JCAM, 2001.

[24] Ford N.J. and Simpson A.C., The numerical solution of fractional differential equations: Speed versus accuracy, Numerical Algorithms, 26, 333-346, 2001.

[25] Gorenflo R. and Luchko Y., *The Initial Value Problem For Some Fractional differential Equations*, Acta Mathematica Vietnamica, 24, 207-233, 1999.

[26] Hadamard J., *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Yale University Press, 1923.

[27] Hairer E., Norsett S.P. and Wanner G., *Solving Ordinary Differential Equations I*, Springer-Verlag, 1991.

[28] Hairer E., Lubich Ch. and Schlichte M., *Fast numerical solutions of weakly singular Volterra integral equations*, Journal of Computational and Applied Mathematics 23, 87-98, 1988.

[29] Henrici P., *Fast Fourier Methods in Computational Complex Analysis*, SIAM Review 21, 1979.

[30] Hilfer R., *Applications of Fractional Calculus in Physics*, World Scientific, 2000.

[31] Ince E.L. *Ordinary Differential Equations*, Dover Publications New York, 1956.

[32] Iserles A., *A First Course in the Numerical Analysis of Differential Equations*, Cambridge Texts in Applied Mathematics, 1997.

[33] Keinert F. and Makroglou A., *On some recent numerical methods for solving linear weak singular Volterra equations of the first kind; a comparative survey*, HERCMA, Athens, 2001.

[34] Kolmogorov A. N. and Fomin S. V., *Fundamentals of the Theory of Functions and Functional Analysis*, Nauka, Moscow, 1968.

[35] Krommer A.R. and Ueberhuber C.W., *Computational Integration*, SIAM, 1998.

[36] Lambert J.D., *Numerical Methods for Ordinary Differential Equations*, Wiley, 1991.

[37] Linz P., *Analytical and Numerical Methods for Volterra Equations*, SIAM, 1985.

[38] Liouville J., *Memoire sur quelques questions de geometrie et de mecanique, et sur un nouveau genre de calcul pour resoudre ces questions*, Journal l'Ecole Royal Polytechnique, 13, Section 21, 1832.

[39] Locket F. J.. *Non-linear Viscoelastic Solids* Academic Press, 1972.

[40] Lubich C., *Discretised Fractional Calculus*, SIAM Journal of Mathematical Analysis, Volume 17, No. 3, 704-719, 1986.

[41] Lubich C., *Fractional Linear Multistep Methods for Abel-Volterra Integral Equations of the First Kind*, IMA Journal of Numerical Analysis, 97-106, 7.

[42] Lubich C., *Convolution Quadrature and Discretised Operational Calculus. I*, Numerische Mathematik, 52, 129-145, 1988.

[43] Lubich C., *Convolution Quadrature and Discretised Operational Calculus. II*, Numerische Mathematik, 52, 413-425, 1988.

[44] Mainardi F., *Fractals and Fractional Calculus*, Springer-Verlag, 1997.

[45] Miller K.S. and Ross B., *An Introduction to the Fractional Calculus and Fractional Differential Equations*, John Wiley, 1993.

[46] Oldham K.B. and Spanier J., *The Fractional Calculus*, Academic Press, 1974.

[47] Pipkin A.C., *Lectures on Viscoelasticity Theory*, George Allen & Unwin Ltd., 1972.

[48] Podlubny I., *Fractional Differential Equations*, Academic Press, 1999.

[49] Podlubny I., *Numerical Solution of Ordinary Fractional Differential Equations by the Fractional Difference Method*, Proceedings of the Second International Conference in Difference Equations, Gordon and Breach Scientific Publishers, 507-515, 1997.

[50] Ralston A. and Rabinowitz P, *A First Course In Numerical Analysis*, McGraw-Hill, 1978.

[51] Riemann B., *Versuch einer allgemeinen Auffassung der Integration und Differentiation*, Gesammelte Mathematische Werke und Wissenschaftlicher, Dover, 1953

[52] Robinson A., *Non-standard Analysis*, North Holland, 1966.

[53] Samko S.G., Kilbas A.A., Marichev O.I., *Fractional Integrals and Derivatives Theory and Applications*, Gordon and Breach, 1993

[54] Scott-Blair G.W., *Survey of General and Applied Reology*, Pitman, 1949.

[55] Simpson A.C., *Linear Fractional Order Differential Equations and their Solution*, M.Sc. Thesis, University College Chester, 1996.

[56] Sloan I.H. and Thomee V., *Time Discretisation of an Integro-Differential Equation of Parabolic Type*, SIAM Journal on Numerical Analysis, 23, 1052-1061, 1986.