Modelling of Accretion Discs with Smoothed Particle Hydrodynamics

Katherine Joan Manson
December 2004

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University.

MATHEMATICAL SCIENCES INSTITUTE
AUSTRALIAN NATIONAL UNIVERSITY
I declare that this thesis contains my own original research. This work has not been submitted previously in any form for any degree at another university. Any information from published and unpublished work has been referenced and acknowledged in the body of the text and listed in the bibliography.

[Signature]

K. Manson
Acknowledgements

It's been a long four years, and I never would have made it this far without the assistance and support of many people.

Firstly, of course, I must thank my principal supervisor, Prof. Wickramasinghe. Dayal's enthusiasm and interest in the many directions this project could have taken has been unmatched. His patience with slow results, and no results was incredible.

Dr. James Murray has also been a tremendous help. James was there at the beginning teaching me how to code a linked list, and at the end reading drafts, and all the points in between. Thank you for being such a good friend and patient and supportive advisor.

Thanks also to Steve Roberts for teaching me the finer points of numerical computation, and to Rainer Wehrse, for teaching me all about radiative transfer. I only hope that the radiative transfer code we wrote will be in use soon.

Thanks to the boys at ANU for all of the toasted sandwiches and Friday afternoon beers, and for the long list of possible thesis topics, should mathematics or astrophysics ever get too boring. Thanks to everyone at Swinburne for showing me such a good time in Melbourne, I decided to move back permanently. Thank you for all of the coffee breaks, and for giving me work when I was running out of money. I apologise for the hammering the cluster has taken in the past few weeks. I'm done now, I promise.

Thank you so much to David Barnes and Rachel Webster at Melbourne Uni for allowing me the time and space to finish the evil beast and for being so very patient and supportive whilst I was writing up. Thanks to the other students for welcoming me and making me feel so at home. Thanks for the thesis haiku and games of pool. Thanks especially to Jamie for pointing out at the last minute that maybe I should check that the coriolis force was right.

My family have always been so supportive of me. Thanks to my parents and sisters for the many long distance phone calls and emails. Mum – I'm glad I finished in time for you to read it.

And finally, the most thanks of all go to Bridget. I couldn't have finished
this without you. Thank you so much for everything.
Abstract

This thesis begins by examining the Lagrangian method of solving fluid equations known as Smoothed Particle Hydrodynamics (SPH) with particular emphasis on its application to the modelling of accretion discs. Several methods of improving the efficiency of SPH calculations are discussed, including varying the smoothing length in both space and time and methods for determining which particles are neighbours.

The importance of the algorithm used for stepping forward in time is well known – it influences not only the efficiency of the calculations but the accuracy and stability as well. Time stepping algorithms are discussed in depth, in particular the method proposed by Strang [72]. Its accuracy is determined, and a method for choosing the timestep size which ensures stability and accuracy is proposed.

The code written for this thesis is then tested. It is found to perform well when modelling a Riemann shock tube and a viscously spreading ring. The code tests of Imaeda and Inutsuka [22] are repeated, and density, velocity and pressure profiles are followed for more than five orbital periods – nearly twice as long as Imaeda and Inutsuka – with no discernable errors found.

The code is then applied to modelling accretion discs. First, discs around isolated one solar mass objects are examined. Analytic solutions are obtained under a number of assumptions, and these equations are modelled. The modelling of the inner boundary is examined and a new method proposed which gives much more accurate results in this region. The energy equation in the disc is examined in depth, and simulations are run to determine the effects of the cooling term on the disc.

Next, accretion discs in cataclysmic variables are modelled. The effect of the mass ratio of the binary is examined, and simulations are presented which demonstrate disc precession.

Finally, the system TV Columbae is studied in depth. TV Col is known to have a mass ratio of approximately 0.7 – 0.96, which should theoretically preclude disc precession, however this system is observed to have both positive and negative superhumps [64]. The inner boundary in TV Col is
discussed, with emphasis on what effect it may have on precession in the disc. Simulations are run to examine these effects.
## Contents

1 Introduction 1

### I SPH Theory and Code 5

2 SPH Basics 7  
2.1 Historical Overview 7  
2.2 Theoretical Overview 8  
2.3 The Equations of Fluid Dynamics 10  
  2.3.1 Equation of Continuity 11  
  2.3.2 Equation of Acceleration 12  
  2.3.3 The Energy Equation 14  
2.4 The Kernel 16  
2.5 The Gravitational Force 16  
2.6 Viscosity 18  
2.7 Viscous Heating 20  
2.8 Scaling the Equations 21  
2.9 Allowing h to vary 26  
2.10 Neighbour searching 27  
2.11 Summary 30

3 Numerical Methods 31  
3.1 Time stepping 31  
3.2 Strang Splitting 33  
3.3 Numerical Stability and Error Control 37  
3.4 Summary 40

4 Testing the Code 41  
4.1 The Shock Tube test 41  
4.2 The ring test 43  
4.3 A Rebuttal 47
4.4 Summary ........................................ 48

II Modelling Accretion Discs 61

5 Discs Around Isolated Objects 63
  5.1 Steady Thin Disc Theory ......................... 64
  5.2 The Inner Boundary ................................ 73
  5.3 The Energy Equation ............................... 76
  5.4 Disc Simulations .................................. 79
  5.5 Summary ........................................... 84

6 Cataclysmic Variable Discs 97
  6.1 The Basic Model .................................. 97
  6.2 Cataclysmic Variables .............................. 100
  6.3 SPH Simulations of Cataclysmic Variables .......... 103
  6.4 Simulations ....................................... 104
  6.5 Varying the mass ratio ............................ 106
  6.6 The Viscosity Parameters .......................... 123
  6.7 Summary .......................................... 124

7 TV Columbae – A Hunt for Precession 131
  7.1 TV Col ............................................ 131
  7.2 Simulation Details .................................. 133
  7.3 Results ........................................... 135
  7.4 Summary .......................................... 136

8 Conclusions and Further Work 151
## List of Figures

4.1 The shock tube test: h constant ........................................... 49
4.2 The shock tube test: h variable, updated using equation (2.16) 50
4.3 The shock tube test: h variable, updated using equation (2.15) 51
4.4 The initial setup for the ring test ......................................... 52
4.5 Results for the ring test ..................................................... 52
4.6 Results for the ring test: longer evolution ............................. 53
4.7 The ring test with variable smoothing length .......................... 53
4.8 The ring test with variable smoothing length: longer evolution 54
4.9 Angular momentum - constant h ......................................... 54
4.10 Angular momentum - variable h ......................................... 55
4.11 The first simulation of Imaeda and Inutsuka: radial density profile .............................................................. 56
4.12 The first simulation of Imaeda and Inutsuka: particle positions 57
4.13 The second simulation of Imaeda and Inutsuka: radial density profile .............................................................. 58

5.1 Normal inner boundary condition. x axis is radius, y axis is density. .............................................................. 74
5.2 Buffer particle inner boundary condition. x axis is radius, y axis is density .............................................................. 75
5.3 Timesteps used in the Thick2 and Thin3 simulations ................. 83
5.4 Simulation Thick1: Energy profiles ..................................... 85
5.5 Simulation Thick1: Density profiles ..................................... 86
5.6 Simulation Thick2: Energy profiles ..................................... 87
5.7 Simulation Thick2: Density profiles ..................................... 88
5.8 Simulation Thick3: Energy profiles ..................................... 89
5.9 Simulation Thick3: Density profiles ..................................... 90
5.10 Simulation Thin1: Energy profiles ..................................... 91
5.11 Simulation Thin1: Density profiles ..................................... 92
5.12 Simulation Thin2: Energy profiles ..................................... 93
5.13 Simulation Thin2: Density profiles ..................................... 94
5.14 Simulation Thin3: Energy profiles ..................................... 95
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.15</td>
<td>Simulation Thin3: Density profiles</td>
<td>96</td>
</tr>
<tr>
<td>6.1</td>
<td>Roche Equipotentials for $m_1 = 2m_2$</td>
<td>98</td>
</tr>
<tr>
<td>6.2</td>
<td>Roche Potential Surface for $m_1 = 2m_2$</td>
<td>99</td>
</tr>
<tr>
<td>6.3</td>
<td>The family tree of cataclysmic variables</td>
<td>101</td>
</tr>
<tr>
<td>6.4</td>
<td>Simulation A: building the disc with $q=0.2$</td>
<td>108</td>
</tr>
<tr>
<td>6.5</td>
<td>Simulation B: building the disc with $q=0.8$</td>
<td>109</td>
</tr>
<tr>
<td>6.6</td>
<td>The longer term evolution of the disc with $q=0.2$</td>
<td>111</td>
</tr>
<tr>
<td>6.7</td>
<td>The longer term evolution of the disc with $q=0.8$</td>
<td>112</td>
</tr>
<tr>
<td>6.8</td>
<td>Relevant disc radii with discs in cataclysmic variables.</td>
<td>116</td>
</tr>
<tr>
<td>6.9</td>
<td>Precessing Disc</td>
<td>117</td>
</tr>
<tr>
<td>6.10</td>
<td>Precessing Disc</td>
<td>118</td>
</tr>
<tr>
<td>6.11</td>
<td>Simulation B: Density</td>
<td>119</td>
</tr>
<tr>
<td>6.12</td>
<td>Simulation B: Dissipation</td>
<td>120</td>
</tr>
<tr>
<td>6.13</td>
<td>Lombic periodogram of the dissipation from the precessing disc. x axis is frequency, y axis is spectral power.</td>
<td>121</td>
</tr>
<tr>
<td>6.14</td>
<td>Lombic periodogram of the dissipation from the non-precessing disc. x axis is frequency, y axis is spectral power.</td>
<td>122</td>
</tr>
<tr>
<td>6.15</td>
<td>Building the disc with $q=0.2$, $\alpha = 10$ and $\beta = 0$.</td>
<td>125</td>
</tr>
<tr>
<td>6.16</td>
<td>The further evolution of the disc with $q=0.2$, $\alpha = 10$ and $\beta = 0.126$</td>
<td>126</td>
</tr>
<tr>
<td>6.17</td>
<td>Precessing Disc</td>
<td>127</td>
</tr>
<tr>
<td>6.18</td>
<td>Precessing Disc: Dissipation</td>
<td>128</td>
</tr>
<tr>
<td>6.19</td>
<td>Lombic periodogram of the dissipation from the disc with $q=0.2$, $\alpha = 10$ and $\beta = 0$. x axis is frequency, y axis is spectral power.</td>
<td>129</td>
</tr>
<tr>
<td>7.1</td>
<td>Power spectrum, Simulation A. x axis is frequency, y axis is spectral power</td>
<td>138</td>
</tr>
<tr>
<td>7.2</td>
<td>Power spectrum, Simulation B. x axis is frequency, y axis is spectral power</td>
<td>139</td>
</tr>
<tr>
<td>7.3</td>
<td>Power spectrum, Simulation C. x axis is frequency, y axis is spectral power</td>
<td>140</td>
</tr>
<tr>
<td>7.4</td>
<td>Power spectrum, Simulation D. x axis is frequency, y axis is spectral power</td>
<td>141</td>
</tr>
<tr>
<td>7.5</td>
<td>Power spectrum, Simulation E. x axis is frequency, y axis is spectral power</td>
<td>142</td>
</tr>
<tr>
<td>7.6</td>
<td>Power spectrum, Simulation F. x axis is frequency, y axis is spectral power</td>
<td>143</td>
</tr>
<tr>
<td>7.7</td>
<td>Simulation A: Density. xy plane coloured according to density.</td>
<td>144</td>
</tr>
<tr>
<td>7.8</td>
<td>Simulation A: Dissipation. xy plane coloured according to dissipation.</td>
<td>145</td>
</tr>
</tbody>
</table>
7.9 Simulation C: Density. xy plane coloured according to density. 146
7.10 Simulation C: Dissipation. xy plane coloured according to dissipation. 147
7.11 Simulation F: Density. xy plane coloured according to density. 148
7.12 Simulation F: Dissipation. xy plane coloured according to dissipation. 149
Chapter 1

Introduction

Until the discovery of U Geminorum in 1855, the only observations of cataclysmic variables were those of a handful of classical novae. It was discovered by John Hind, who was searching for minor planets [76]. When U Gem returned to maximum light several months after its initial discovery, it proved to be no ordinary nova. Together with SS Cygni, discovered in 1896 by Louisa D. Wells, it constituted a new class of star. This new class was originally termed ‘subnovae’ because of the much smaller amplitude of outburst when compared to classical novae. They are now known as ‘dwarf novae’.

In 1940 when Joy [23] observed the dwarf nova RU Peg at minimum light, he discovered that it had both an absorption and emission line spectrum, which suggested the presence of two stars. In 1954, Walker [75] found that DQ Her was an eclipsing binary, and in 1956 Joy determined [24] that SS Cyg was also a binary star. In order to test the theory that all such stars may be binaries, Kraft [25] presented a spectroscopic study of six known U Gem stars, finding strong evidence that four of them were binaries, and no evidence to contradict the theory. Further, he presented evidence to suggest that the secondary star might fill its Roche lobe. He postulated that a stream of material “spills out from the inner Lagrangian point and forms a ring or disc around the blue companion”.

The current model of a cataclysmic variable consists of a white dwarf accreting material from a low-mass companion and, in general, it is the process
of accretion that makes cataclysmic variables so interesting. Accretion is the most effective energy release mechanism known, producing more than twenty times the amount of energy produced by converting hydrogen to helium. In many cataclysmic variables, the material being accreted passes first through a disc of matter surrounding the white dwarf, and it is this disc which is the primary concern of this thesis.

Since Kraft [25] and Crawford and Kraft [9] first proposed that accretion might be behind the outburst behaviour of AE Aqr and the U Gem stars, the accretion disc model has been used to explain many of the rich variety of phenomena found in cataclysmic variables. Changes in the viscosity or mass transfer rate in accretion discs may account for dwarf novae outbursts; a thermonuclear runaway in the material built up on the surface of the white dwarf is responsible for the eruptions of classical novae; a non-symmetric, precessing accretion disc explains the presence of superhumps; and so on.

Early models of accretion discs were made based on many assumptions, such as requiring the disc to be axisymmetric, physically thin and optically thick; that material in the disc follows nearly Keplerian orbits; that the kinematic viscosity and opacity were constant throughout the disc. When it was introduced in 1977, the Smoothed Particle Hydrodynamics method ([36] and [11]) allowed many of these assumptions to be dropped. Since then, it has been built upon and used to simulate and investigate many interesting aspects of accretion discs in many different situations.

This thesis is divided into two parts. The first part focusses on the theory, methods and algorithms involved in using smoothed particle hydrodynamics. To begin with, the theory behind SPH is introduced and the equations used throughout the thesis examined. Two of the more common ways of improving the efficiency of an SPH simulation are also presented in chapter 2.

SPH provides a method of determining the derivatives of a function at a set of disordered points, but these derivatives must then be stepped forward in time. The problem of numerical integration is addressed in chapter 3. This chapter also discusses the numerical stability and accuracy of such an integration, along with a method for controlling error.

The final chapter in part one presents the results of several tests of the
SPH code. These tests are designed to ensure that viscosity is modelled properly, that shocks are resolved, and that those situations which have analytic solutions are accurately modelled. They are a part of a larger, more comprehensive, suite of tests which ensure that any results produced by the code are reliable.

The second part of this thesis presents applications of the code developed in part I – specifically, it concentrates on models of accretion discs. The first model presented is of a thin disc accreting onto an isolated object. The theory of thin discs is derived using the Shakura–Sunyaev viscosity parametrisation. Problems with modelling the inner boundary of an accretion disc are discussed and a solution proposed. Finally, the energy equation in the disc is examined, and the effects of changing the balance between heating and cooling modelled.

The focus shifts in chapter 6 to accretion discs in cataclysmic variables. The effect of the mass ratio of the binary on the morphology of the disc is investigated, and the disc precession thought to be behind the phenomenon of superhumps is studied.

In chapter 7, a specific system is studied. TV Columbae is an intermediate polar whose mass ratio is believed to be well above the theoretic limit for disc precession, yet the system shows both positive and negative superhumps. In this chapter, I will present a possible explanation for the superhumps seen in TV Col. A number of simulations designed to test this theory are carried out, however no precession is found.
Part I

SPH Theory and Code
Chapter 2

SPH Basics

In this chapter, I will introduce the smoothed particle hydrodynamics method. I will discuss the central theory of SPH and introduce the hydrodynamical equations which I will be modelling.

2.1 Historical Overview

There are many methods for modelling the flow of a fluid. Some require that equations be calculated on a grid, either static or moving; some involve calculating the equations over a set of moving, interacting particles; some are a combination of both a grid and a set of particles (for example, the particle in cell method - see [14] and references therein). The method I have chosen to investigate this problem is known as smoothed particle hydrodynamics, or SPH. It was first introduced in 1977 by Gingold and Monaghan [11], who tested it on non-axisymmetric polytropes, and Lucy [36], who tested the fission hypothesis of binary evolution. Both of these tests involved non-symmetric, three dimensional problems - problems which were a great challenge to the standard grid-based methods in use at the time.

Since then, the method has been used to solve a wide variety of different problems, from basic gas dynamic problems, to rocky impact problems [61],
to cosmological simulations [17]. Improvements in computer memory and speed have seen the size of the calculations grow from an initial 80 – 200 particles to more than a million. Different neighbour searching schemes have been described, with improved efficiency for different problems. Other extensions have included using variable mass particles, allowing particles to split or join together, and allowing the size of the interpolation kernel to vary over time and/or space. The extensions which I have used in this thesis will be described in detail later in this chapter.

2.2 Theoretical Overview

At the centre of the SPH method is the summation interpolant. This allows us to interpolate an equation from a set of disordered points. These points are often referred to as particles, and this interpretation is useful, since it allows us to intuitively determine interactions between them using basic Newtonian mechanics. More properly though, the points should be thought of as moving interpolation nodes.

Suppose we have a set of points, \( r_1, r_2, \ldots, r_N \), with some property defined by a function \( A \), which could be density, pressure, velocity, and so on.

By using the Dirac delta function, \( \delta(r) \), which has the properties

\[
\delta(x) = 0 \quad \text{if} \quad x \neq 0 \\
\int_{-\epsilon}^{\epsilon} \delta(x) \, dx = 1, \epsilon > 0 \\
\int_{-\infty}^{\infty} F(x) \delta(x - a) \, dx = F(a),
\]

we can write the identity

\[
A(r) = \int A(r') \delta(r - r') \, dr'
\]

where the integration is over the domain, and \( dr' \) is an element of the domain (e.g. \( dx \) in one dimension, or \( dx dy dz \) in three dimensions).
2.2. THEORETICAL OVERVIEW

If we replace the delta function, $\delta(r)$ by a kernel, $W(r, h)$, where

\[
W(r, h) \rightarrow \delta(r) \text{ as } h \rightarrow 0
\]

and \( \int W(r, h)dr = 1 \)

we can then define an integral interpolant:

\[
A_I(r) = \int A(r')W(r - r', h)dr'.
\]

However, since we only know the value of the function $A$ at the points $r_1, r_2, \ldots, r_N$, the integral cannot be evaluated. If we first rewrite the integral as

\[
A_I(r) = \int \frac{A(r')}{\rho(r')}W(r - r', h)\rho dr'
\]

where $\rho$ represents density, we can define the summation interpolant which approximates the integral as

\[
A_S(r) = \sum_b m_b \frac{A_b}{\rho_b} W(r - r_b, h).
\] (2.1)

In doing so, we have recognised that $\rho dr'$ is simply an element of mass. Here, the subscript $b$ indicates a particle label, and the summation is over all of the particles. $A_b$ represents the value of the function $A$ at particle $b$, $\rho_b$ the density at particle $b$ and so on.

If we ensure that the kernel, $W$, is differentiable, then it is simple to differentiate the summation interpolant $A_S$. For instance, the gradient of $A_S$ becomes:

\[
\nabla A_S(r_a) = \sum_b m_b \frac{A_b}{\rho_b} \nabla_a W_{ab}
\] (2.2)

The notation $\nabla_a W_{ab}$ indicates that we are taking the gradient of $W(r_a - r_b)$ with respect to particle $a$. 
2.3 The Equations of Fluid Dynamics

The equations which describe a three-dimensional fluid flow are the equations of continuity, acceleration, and energy. These may be written (without any viscosity, heating or cooling terms or external forces):

\[
\text{continuity: } \frac{d\rho}{dt} = -\rho \nabla \cdot \mathbf{v} \tag{2.3}
\]
\[
\text{acceleration: } \frac{d\mathbf{v}}{dt} = -\frac{1}{\rho} \nabla P \tag{2.4}
\]
\[
\text{energy: } \frac{dU}{dt} = -\frac{P}{\rho} \nabla \cdot \mathbf{v} \tag{2.5}
\]

Here, \(\mathbf{v}\), \(\rho\), \(P\) and \(U\) represent the velocity, density, pressure, and thermal energy of the fluid per unit mass, respectively. The derivatives used are Lagrangian time derivatives (the rate of change following the motion of the fluid), that is:

\[
\frac{df}{dt} = \frac{\partial f}{\partial t} + (\mathbf{v} \cdot \nabla) f
\]

This gives us three equations in four variables, and we need an equation of state to close the system. Other than degenerate gases in white dwarfs and neutron stars, and the cores of normal stars, an astrophysical gas is well described by the ideal gas law:

\[
P = \frac{\rho kT}{\mu m_H}
\]

Here \(T\) is the temperature of the gas, \(k\) is Boltzmann's constant, \(m_H\) is the mass of the hydrogen atom and \(\mu\) is the average mass per particle of the gas measured in terms of \(m_H\). Although it may seem that we have introduced another variable – \(T\) – there is a simple relation between \(T\) and \(U\), the thermal energy of the gas, which comes from the equipartition theorem of kinetic theory. This relation says that every degree of freedom of each gas particle is assigned an energy \(\frac{1}{2}kT\). For a monatomic gas, there are three
degrees of freedom for each particle, and so:

\[ U = \frac{3kT}{24\mu m_H} \quad (2.6) \]

While the gases with which we are concerned will not necessarily be monatomic, this is a good approximation.

### 2.3.1 Equation of Continuity

The simplest way of deriving the SPH form of the equation of continuity is to begin with the equation for the density, then find the time derivative of that. Using equation (2.1), we have that

\[ \rho_a = \sum_b m_b W_{ab} \quad (2.7) \]

and so,

\[
\frac{d\rho_a}{dt} = \frac{d}{dt} \sum_b m_b W_{ab} \\
= \sum_b m_b \frac{dW_{ab}}{dt} \\
= \sum_b m_b \nabla_a W_{ab} \cdot \frac{dr_{ab}}{dt}
\]

However,

\[
\frac{d}{dt} r_{ab} = \frac{d}{dt} (r_a - r_b) \\
= v_a - v_b \\
= v_{ab}
\]

so

\[
\frac{d\rho_a}{dt} = \sum_b m_b v_{ab} \cdot \nabla_a W_{ab}. \quad (2.8)
\]
We can now decode this expression by writing each of the SPH terms in their continuous equivalents:

\[
\sum_b m_b \mathbf{v}_{ab} \cdot \nabla_a W_{ab} = \mathbf{v}_a \cdot \sum_b m_b \nabla_a W_{ab} - \sum_b m_b \mathbf{v}_b \nabla_a W_{ab}
\]

\[
= \mathbf{v}_a (\nabla \rho)_a - (\nabla \cdot (\rho \mathbf{v}))_a
\]

\[
= -(\rho \nabla \cdot \mathbf{v})_a
\]

where we have used a vector identity in the last step. In doing this, we see that the SPH equation for the rate of change of density reduces to the standard equation of continuity for fluids, equation (2.3).

When modelling a fluid with SPH, one has the choice of using either equation (2.7) or equation (2.8) to calculate the density. The advantage of using (2.8) is that all the derivatives may be calculated in one pass over the computational field (see section 2.10 for more detail), therefore making the computation more efficient. However, using equation (2.7) is computationally far more robust and stable. In this work, I have chosen to use (2.7) to calculate density.

2.3.2 Equation of Acceleration

We begin with finding an SPH expression for the gradient of the pressure. Using the equation (2.2) as a guide, we can write:

\[
(\nabla P)_a = \sum_b m_b \frac{P_b}{\rho_b} \nabla_a W_{ab}
\]

Then, the SPH approximation to equation (2.4) is:

\[
\frac{d\mathbf{v}_a}{dt} = -\frac{1}{\rho_a} \sum_b m_b \frac{P_b}{\rho_b} \nabla_a W_{ab}
\]
This equation, however, does not conserve angular momentum. Instead, we may write:

\[
\frac{\nabla P}{\rho} = \nabla \left( \frac{P}{\rho} \right) + \frac{P}{\rho^2} \nabla \rho
\]

we have:

\[
\nabla \left( \frac{P}{\rho} \right) = \sum_b m_b \frac{P_b}{\rho_b^2} \nabla_a W_{ab}
\]

\[
\frac{P}{\rho^2} \nabla \rho_a = \frac{P_a}{\rho_a^2} \sum_b m_b \nabla_a W_{ab}
\]

Summing these terms, we have that:

\[
\frac{\nabla P}{\rho} = \sum_b m_b \frac{P_b}{\rho_b^2} \nabla_a W_{ab} + \frac{P_a}{\rho_a^2} \sum_b m_b \nabla_a W_{ab}
\]

\[
= \sum_b m_b \frac{P_b}{\rho_b^2} \nabla_a W_{ab} + \sum_b m_b \frac{P_a}{\rho_a^2} \nabla_a W_{ab}
\]

\[
= \sum_b m_b \left( \frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} \right) \nabla_a W_{ab}
\]

That is,

\[
\frac{dv_a}{dt} = - \sum_b m_b \left( \frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} \right) \nabla_a W_{ab}
\]

This form of the acceleration equation is symmetric, that is, the force which particle \( a \) exerts on particle \( b \) is the same as that which particle \( b \) exerts on particle \( a \), and it conserves linear momentum exactly. If the force is also along the line of centre of the two particles then angular momentum will also be conserved exactly. Whether or not this condition holds depends on the kernel used. The kernel used throughout this thesis (see section 2.4) conserves angular momentum.
2.3.3 The Energy Equation

In the derivation of the SPH form of the acceleration equation, we were able to use a vector identity to arrive at a symmetric form. Similarly, we may do that here, though the derivation is a little more complex.

We begin by rewriting the fluid equation:

\[
\frac{dU}{dt} = -\frac{P}{\rho} \nabla \cdot \mathbf{v}
\]

\[
= -\frac{P}{\rho^2} (\rho \nabla \cdot \mathbf{v})
\]

\[
= -\frac{P}{\rho^2} [\nabla \cdot (\rho \mathbf{v}) - \mathbf{v} \cdot \nabla \rho]
\]

By using equations (2.1) and (2.2), we can get

\[
\nabla \cdot (\rho \mathbf{v}) = \sum_b m_b v_b \cdot \nabla_a W_{ab}
\]

and

\[
\mathbf{v} \cdot \nabla \rho = v_a \sum_b m_b \nabla_a W_{ab}
\]

and so:

\[
\frac{dU}{dt} = \frac{P_a}{\rho_a} \sum_b m_b (v_a - v_b) \nabla_a W_{ab}
\]

(2.9)

We now write equation (2.5) in a different way:

\[
\frac{dU}{dt} = -\frac{P}{\rho} \nabla \cdot \mathbf{v}
\]

\[
= -\nabla \left( \frac{P \mathbf{v}}{\rho} \right) + \mathbf{v} \cdot \nabla \left( \frac{P}{\rho} \right)
\]
Again using (2.1) and (2.2), we get:

\[-\nabla \left( \frac{Pv}{\rho} \right) = -\nabla \cdot \sum_b m_b \frac{P_b}{\rho_b^2} v_b W_{ab} = -\sum_b m_b \frac{P_b}{\rho_b^2} v_b \cdot \nabla W_{ab}\]

and

\[v \cdot \nabla \left( \frac{P}{\rho} \right) = v_a \cdot \sum_b m_b \frac{P_b}{\rho_b^2} \nabla W_{ab}\]

and so:

\[
\frac{dU}{dt} = \sum_b m_b \frac{P_b}{\rho_b^2} (v_a - v_b) \cdot \nabla W_{ab} \tag{2.10}
\]

We now take an average of the two equations, (2.9) and (2.10) to get a symmetric form for the energy equation.

\[
\frac{dU}{dt} = \frac{1}{2} \left( \frac{dU}{dt} + \frac{dU}{dt} \right) = \frac{1}{2} \left( \frac{P_a}{\rho_a^2} \sum_b m_b (v_a - v_b) \nabla W_{ab} + \sum_b m_b \frac{P_b}{\rho_b^2} (v_a - v_b) \cdot \nabla W_{ab} \right) = \frac{1}{2} \sum_b m_b \left( \frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} \right) v_{ab} \cdot \nabla W_{ab}.
\]

Although this form of the energy equation conserves total energy, Benz [3] notes that it can largely overestimate the energy derivative, resulting in negative energies. In practice, I have not found any evidence of this.
2.4 The Kernel

In this thesis, I have used the spline function interpolant first introduced by Monaghan and Lattanzio in 1985 [45]. It takes the form:

\[ W(r, h) = \frac{\sigma}{h^D} \begin{cases} 
1 - \frac{3}{2}x^2 + \frac{3}{4}x^3 & 0 \leq x \leq 1 \\
\frac{1}{4} (2 - x)^3 & 1 \leq x \leq 2 \\
0 & \text{else}
\end{cases} \]

where \( x = r/h \), \( D \) is the number of dimensions, and \( \sigma \) is a normalisation constant equal to 2/3, 10/7\( \pi \) or 1/\( \pi \) in 1, 2 and 3 dimensions respectively.

This kernel has compact support (that is, the kernel is zero outside a certain radius), which means that particles further apart than the support radius (2\( h \)) will not influence one another. Its first derivatives are smooth, which means that the SPH method can be used to estimate function derivatives, and the second derivative is smooth, which means that the estimations are not sensitive to particle disorder.

2.5 The Gravitational Force

Aside from the force due to pressure gradients, each particle will experience a force due to gravity. This will of course depend on the situation, and there are two cases I will cover here. In both cases, the gravitational force is an external force; the particles are not self-gravitating.

The first of these cases is that of a single point object. This would be the case for particles orbiting a single star, and it is this case which we use for the ring test.

With the central object placed at the origin, the force per unit mass on a particle is given by

\[ F = -\frac{GMr}{r^3}. \]

(2.11)

Here \( G \) is the gravitational constant, \( M \) is the mass of the central object and
\( \mathbf{r} \) is the radius vector.

The second case is that of a particle in a binary system. In this case, the particle will experience a gravitational force due to each of the stars, as well as an additional force due to the rotation of the binary. In order to calculate the potential experienced by a test particle in such a system, there are a number of assumptions we must make. Firstly, we assume that each of the stars may be represented by a point mass. Secondly, we assume that these stars are in circular, Keplerian orbits about the system’s centre of mass. Finally, we must assume that the mass of the test particle is negligible compared to that of the stars, so that it does not perturb their orbits, nor does it transfer angular momentum to either of the stars.

If we place the primary object at the origin, and align the axes so that the secondary object sits on the x-axis (noting that the frame of reference is rotating with the binary), then the potential energy of a particle is given by the Roche potential:

\[
\phi = \frac{GM_1}{r_1} + \frac{GM_2}{r_2} + \frac{\omega^2}{2} \left[ (x - x_{cm})^2 + y^2 \right]
\]

where \( G \) is again the gravitational constant, \( M_1 \) is the mass of the primary, \( M_2 \) the mass of the secondary, \( r_1 \) the distance from the particle to the primary, \( r_2 \) the distance from the particle to the secondary, \( \omega \) is the size of the angular velocity of the system and \( x_{cm} \) is the (x-coordinate) position of the centre of mass.

In order to find the force per unit mass on the particle, we take the gradient of this potential and, because we are operating in a rotating frame of reference, add the Coriolis term:

\[
\mathbf{F} = \nabla \phi - 2\Omega \times \mathbf{v}.
\]

In this coordinate system, the angular velocity lies along the z axis, and has
size $\omega$, so

\[
F_x = -\frac{GM_1 x}{r_1^3} + \frac{GM_2 (a - x)}{r_2^3} + \omega^2 (x - cm) + 2\omega v_y
\]

\[
F_y = -\frac{GM_1 y}{r_1^3} - \frac{GM_2 y}{r_2^3} + \omega^2 y - 2\omega v_x
\]

\[
F_z = -\frac{GM_1 z}{r_1^3} - \frac{GM_2 z}{r_2^3}
\]

Here, $a$ is the separation of the primary and secondary.

### 2.6 Viscosity

In a real system, kinetic energy will be dissipated to heat over very short length scales. In the presence of a shock, for example, dissipation occurs over a few mean free path lengths of the molecules in the fluid – a length scale at which even the fluid assumptions of the hydrodynamic equations break down. Unless an additional, artificial viscosity is added to the numerical equations, such dissipation effects are not included. Omission of dissipation is not only unphysical; it also generally leads to numerically unstable simulations [53].

Typically, a macroscopic viscous pressure is added to the equations in order to mimic the microscopic dissipation, and allow kinetic energy to be converted to heat. This artificial viscous pressure can take many forms, but is usually based on two types of viscosity: a bulk viscosity which is linearly proportional to the divergence of the velocity field [36]; and a vonNeumann – Richtmyer viscosity which is proportional to the square of the divergence of the velocity field [74].

Probably the most common form of viscosity used in SPH calculations is that derived by Monaghan and Gingold [47], and uses a combination of the bulk and vonNeumann – Richtmyer viscosities. They define an estimate of the velocity divergence at particle $a$ due to the presence of $b$ by:

\[
\mu_{ab} = \frac{h v_{ab} \cdot \mathbf{r}_{ab}}{r_{ab}^2 + \eta^2}
\]  

(2.12)
where $\eta$ is a small number which is used to keep the divergence from becoming singular when the two interacting particles are very close ($r_{ab} \to 0$). $\eta$ is typically set to be equal to 10% of the smoothing length. The viscosity pressure is then given by

$$\Pi_{ab} = \begin{cases} \frac{-\alpha c_{ab} u_{ab} + \beta \rho^2_{ab}}{\rho_{ab}} & \mathbf{v}_{ab} \cdot \mathbf{r}_{ab} < 0 \\ 0 & \mathbf{v}_{ab} \cdot \mathbf{r}_{ab} > 0 \end{cases}$$

where $\bar{\rho}_{ab} = \frac{1}{2}(\rho_a + \rho_b)$ and $\bar{c}_{ab} = \frac{1}{2}(c_a + c_b)$ are the average values for the density and sound speed respectively.

There are two free parameters here: $\bar{\alpha}$ and $\beta$. ($\bar{\alpha}$ is barred to distinguish it from the Shakura--Sunyaev viscosity parameter $\alpha$.) The linear coefficient, $\bar{\alpha}$, produces both a shear and a bulk viscosity [48], while the quadratic coefficient, $\beta$, is roughly equivalent to the von Neumann--Richtmyer viscosity and is used to handle strong shocks. The choice of values for $\bar{\alpha}$ and $\beta$ will be determined by the conditions being simulated. Based on comparison with shock tube experiments, Lattanzio et al [30] found the values required to damp oscillations behind high Mach number shocks were typically $\bar{\alpha} \approx 1$ and $\beta \approx 2$. These values, however, can lead to excessively high shear and bulk viscosity and can tend to smear out turbulence and flow details. For this reason, when shocks are unimportant, $\beta$ is often set to zero. (See, for example, Meglicki, Wickramasinghe and Bicknell [39]).

The switch present in this form of the viscosity means that the viscosity only acts for compression, and not rarefaction. That is, particles will only experience a viscous force if they are approaching each other. In this thesis I wish to model shear viscosity and so, following the work of Murray [52] and Meglicki, Wickramasinghe and Bicknell [39], I remove the switch and allow particles to interact when they are both approaching and receding, resulting in a viscous pressure of the form:

$$\Pi_{ab} = \begin{cases} \frac{-\alpha c_{ab} u_{ab} + \beta \rho^2_{ab}}{\rho_{ab}} & \mathbf{v}_{ab} \cdot \mathbf{r}_{ab} < 0 \\ \frac{\alpha c_{ab} u_{ab}}{\rho_{ab}} & \mathbf{v}_{ab} \cdot \mathbf{r}_{ab} > 0 \end{cases}$$ (2.13)
This term is added to the acceleration equation

\[
\frac{dv_a}{dt} = - \sum_b m_b \left( \frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} + \Pi_{ab} \right) \nabla_a W_{ab}
\]

It can be shown ([52] and [61]) that this form of \( \Pi_{ab} \) equates to a shear viscosity of magnitude

\[
\nu = \kappa \alpha c_h
\]

plus a bulk viscosity of similar magnitude (which we may safely ignore if the divergence of the velocity field is small). The constant \( \kappa \) relates to the dimensionality of the problem and the kernel used – in two dimensions it is equal to 1/8. The relation can also be written using the Shakura–Sunyaev parameter, \( \alpha \), instead of \( \nu \)

\[
\alpha H = \frac{1}{8} \alpha h
\]  

(2.14)

where \( H \) is the typical scale height of the disc.

### 2.7 Viscous Heating

In its present state, the energy equation does not allow for any cooling, nor does it allow for heating other than compressional heating. The first thing we must add is heating due to viscosity. This is done in the same way as adding viscosity to the acceleration equation, and the form of the viscous term remains the same. That is:

\[
\frac{dU}{dt} = \frac{1}{2} \sum_b m_b \left( \frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} + \Pi_{ab} \right) v_{ab} \cdot \nabla_a W_{ab}
\]

Now we have allowed for viscous and compressional heating.

In this chapter, I will consider only isothermal simulations. That is, \( dU/dt = 0 \). The energy equation is developed alongside the other equations, however indepth consideration of this equation and the role of the cooling term is deferred to chapter 5.
2.8 Scaling the Equations

We wish to scale the equations used in the program to units appropriate to the problem we are studying. The equations we wish to scale are:

\[ \frac{d\rho}{dt} = -\rho \nabla \cdot \mathbf{v} \]  \hspace{1cm} (2.3)

\[ \frac{d\mathbf{v}}{dt} = -\frac{1}{\rho} \nabla P \]  \hspace{1cm} (2.4)

\[ \frac{d\mathbf{U}}{dt} = -\frac{P}{\rho} \nabla \cdot \mathbf{v} \]  \hspace{1cm} (2.5)

\[ F = \frac{GMr}{r^3} \]  \hspace{1cm} (2.11)

and

\[ F_x = \frac{GM_1x}{r_1^3} + \frac{GM_2(a - x)}{r_2^3} + \omega^2(x - cm) + 2\omega v_y \]

\[ F_y = \frac{-GM_1y}{r_1^3} + \frac{-GM_2y}{r_2^3} + \omega^2y - 2\omega v_x \]

\[ F_z = \frac{-GM_1z}{r_1^3} + \frac{GM_2z}{r_2^3} \]

We choose a length \( R \), mass \( M \) and time \( \tau \), and so we have the scaling relations in three dimensions:

\[ \rho = \hat{\rho} \frac{M}{R^3} \]

\[ \mathbf{v} = \hat{\mathbf{v}} \frac{R}{\tau} \]

\[ \nabla = \hat{\nabla} \frac{1}{R} \]

In two dimensions, the density is replaced by the surface density – that is, the density integrated along the z-axis. The surface density is approximated as \( \Sigma = \rho H \) where \( H \) is the typical scale height of the problem. Replacing \( \rho \) with \( \frac{\Sigma}{H} \) (which has identical units to \( \rho \)) in the equations does not change the scaling.
For the first equation:

\[
\frac{d\hat{\rho}}{dt} = -\rho \nabla \cdot \mathbf{v}
\]

\[
\frac{d\hat{\rho}}{d(\tau \hat{R}^3)} = -\hat{\rho} \left( \frac{M}{\hat{R}^3} \right) \left( \frac{1}{\hat{R}} \right) \cdot \left( \mathbf{\hat{v}} \cdot \frac{\hat{R}}{\tau} \right)
\]

\[
\frac{1}{\tau} \frac{d\hat{\rho}}{R^3 \frac{d\tau}{dt}} = -\left( \frac{M}{\hat{R}^3 \tau} \right) \hat{\rho} \nabla \cdot \mathbf{\hat{v}}
\]

\[
\therefore \quad \frac{d\hat{\rho}}{dt} = -\hat{\rho} \nabla \cdot \mathbf{\hat{v}}
\]

For the second equation:

\[
\frac{d\mathbf{v}}{dt} = -\frac{1}{\rho} \nabla P
\]

\[
\frac{d\left( \frac{\hat{\rho} \mathbf{\hat{v}}}{\tau} \right)}{d(\tau \hat{R}^3)} = -\frac{1}{\hat{\rho}} \frac{R^3}{M \hat{R}} \hat{\nabla} (P_{sc} \hat{P})
\]

\[
\frac{R \frac{d\hat{\mathbf{v}}}{\tau^2 \frac{d\tau}{dt}}}{\hat{P}} = \frac{P_{sc} R^2}{M \hat{\rho}} \hat{\nabla} \hat{P}
\]

It’s now time to choose units for \( P_{sc} \) and \( \tau \). For \( \tau \) we use:

\[
\tau^2 = \frac{R^3}{GM}
\]

and we note that this is equivalent to the period of an object in orbit about a mass \( M \), divided by \( 2\pi \). For the pressure, we use:

\[
P_{sc} = \frac{GM^2}{R^4}
\]

a simple dimensional analysis shows that this constant has the same units as pressure. We now substitute these into the above equation:

\[
\frac{RGM \frac{d\hat{\mathbf{v}}}{R^3 \frac{d\tau}{dt}}}{R^3} = -\frac{GM^2 R^2}{M \hat{\rho}} \frac{1}{\hat{\rho}} \hat{\nabla} \hat{P}
\]

\[
\frac{d\hat{\mathbf{v}}}{dt} = -\frac{1}{\hat{\rho}} \hat{\nabla} \hat{P}
\]
A similar procedure holds for the energy equation, and the scaled form is identical to the unscaled form.

In order to scale the gravity equations, we need to find a scaling constant for the force per unit mass. We do this by noting that, in general, $F = ma$, and since we have a force per unit mass, $F = a$.

$$F = a$$
$$= \frac{dv}{dt}$$

$$F_{scF} = \frac{d\left(\frac{\dot{\mathbf{r}}}{r}\right)}{d(t/T)}$$
$$= R\frac{d\hat{\mathbf{v}}}{\tau^2 \frac{dt}{T}}$$
$$= \frac{GM}{R^2} \frac{d\hat{\mathbf{v}}}{dt}$$

This gives us a force scaling constant of:

$$F_{sc} = \frac{GM}{R^2}$$

For the first of the gravity equations, we have:

$$F = \frac{GMr}{r^3}$$

$$F_{scF} = \frac{GM}{\tau^2 R^2}$$

$$\hat{F} = \frac{1}{\tau^2}$$

For the second set of gravity equations, we have:

$$F_x = -\frac{GM_1 x}{r_1^3} + \frac{GM_2 (a - x)}{r_2^3} + \omega^2 (x - cm) + 2\omega v_y$$

$$F_y = -\frac{GM_1 y}{r_1^3} - \frac{GM_2 y}{r_2^3} + \omega^2 y - 2\omega v_x$$

$$F_z = -\frac{GM_1 z}{r_1^3} - \frac{GM_2 z}{r_2^3}$$
I will show the scaling here for the x-direction. The other results are derived in a similar fashion.

\[ F_{sc \hat{x}} = - \frac{G \hat{M}_1 M R \hat{x}}{R^3 \hat{r}_1^3} + \frac{G \hat{M}_2 M R (\hat{a} - \hat{x})}{R^3 \hat{r}_2^3} + \omega_{sc}^2 \hat{\omega}^2 R (\hat{x} - \hat{c} m) + 2 \omega_{sc} \hat{\omega} v_{sc} \hat{v}_y \]

To get the scaling constant for \( \omega \) we note that the units of frequency are inverse time, so we simply use \( 1/\tau \) for \( \omega_{sc} \).

\[ \frac{GM_1 M R \hat{x}}{R^3 \hat{r}_1^3} + \frac{GM_2 M R (\hat{a} - \hat{x})}{R^3 \hat{r}_2^3} + \frac{GM}{R^3} \hat{\omega}^2 R (\hat{x} - \hat{c} m) + 2 \frac{GM}{R^2} \hat{\omega} \hat{v}_y \]

\[ = \frac{GM}{R^2} \left(- \frac{\hat{M}_1 \hat{x}}{\hat{r}_1^3} + \frac{\hat{M}_2 (\hat{a} - \hat{x})}{\hat{r}_2^3} + \hat{\omega}^2 (\hat{x} - \hat{c} m) + 2 \hat{\omega} \hat{v}_y \right) \]

\[ \therefore \hat{F}_x = - \frac{\hat{M}_1 \hat{x}}{\hat{r}_1^3} + \frac{\hat{M}_2 (\hat{a} - \hat{x})}{\hat{r}_2^3} + \hat{\omega}^2 (\hat{x} - \hat{c} m) + 2 \hat{\omega} \hat{v}_y \]

Similarly, we can derive:

\[ \hat{F}_y = - \frac{\hat{M}_1 \hat{y}}{\hat{r}_1^3} - \frac{\hat{M}_2 \hat{y}}{\hat{r}_2^3} + \hat{\omega}^2 \hat{y} - 2 \hat{\omega} \hat{v}_x \]

\[ \hat{F}_z = - \frac{\hat{M}_1 \hat{z}}{\hat{r}_1^3} - \frac{\hat{M}_2 \hat{z}}{\hat{r}_2^3} \]

Finally, we need to scale the equation of state.

\[ P = \frac{\rho k T}{\mu m_H} \]

\[ = \frac{2}{3} \rho U \]

\[ P_{sc \hat{P}} = \frac{2 M - GM \hat{U}}{3 R^3 \rho \hat{r}^3} \]

\[ \hat{P} = \frac{2}{3} \hat{\rho} \hat{U} \]

Note that from this point on, I will drop the hats when referring to scaled quantities. Cases where a quantity is unscaled will be made clear.

To summarise, I include a table of the relevant scaling constants described
above. Note that the base variables $R$ and $M$ remain undefined. These values may be set to whatever is appropriate to the current simulation. For example, when simulating an accretion disc in a cataclysmic variable, I would set $R$ to be the separation between the two stars and $M$ to be their combined mass.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>length</td>
<td>$R$</td>
<td></td>
</tr>
<tr>
<td>mass</td>
<td>$M$</td>
<td></td>
</tr>
<tr>
<td>time</td>
<td>$\tau$</td>
<td>$\sqrt{\frac{R^3}{GM}}$</td>
</tr>
<tr>
<td>density</td>
<td>$\rho_{sc}$</td>
<td>$\frac{M}{R^3}$</td>
</tr>
<tr>
<td>velocity</td>
<td>$v_{sc}$</td>
<td>$\frac{R}{\tau}$</td>
</tr>
<tr>
<td>gradient</td>
<td>$\nabla_{sc}$</td>
<td>$\frac{1}{R}$</td>
</tr>
<tr>
<td>pressure</td>
<td>$P_{sc}$</td>
<td>$\frac{GM^2}{R^4}$</td>
</tr>
<tr>
<td>force</td>
<td>$F_{sc}$</td>
<td>$\frac{GM}{R^2}$</td>
</tr>
</tbody>
</table>

Table 2.1: Summary of scaling constants
2.9 Allowing $h$ to vary

In the derivation of the SPH equations, we have assumed that $h$, the size of the smoothing kernel, is a constant; however this need not be so. We may allow $h$ to vary with both time and position. In doing so, we allow the code to adapt to density conditions, thus maintaining an optimal coverage of the computational domain. In areas where the density decreases, $h$ should increase, so that particles are still interacting, while in areas with increasing density $h$ decreases, which ensures that details are not smoothed out.

There are a number of ways of deriving the equations with $h = h(r, t)$, presented by Monaghan [46], Bicknell [4] and Brookshaw [6], among others. I will not repeat those derivations here, but simply state that so long as the value of the smoothing length used to calculate the kernel is symmetrised between the two contributing particles, the form of the SPH equations remains the same. In order to symmetrise the smoothing length, I simply take an arithmetic mean of the two particles' smoothing lengths.

Since we are allowing $h$ to vary in space and time, we must also decide how we are going to update $h$. Each particle now has its own smoothing length, which it carries in the same way as it has its own pressure, mass and density. Each individual $h$ can then be updated each time step, according to some rule.

Although many forms for changing $h$ have been postulated, the one which I have found to give the best results with the least computational effort is given by

$$h_{\text{new}} = h_{\text{old}} \left( \frac{\rho_{\text{old}}}{\rho_{\text{new}}} \right)^{\frac{1}{d}}$$

(2.15)

where $d$ is the number of dimensions. This form was given by Monaghan [44] and Benz [3]. This method of updating $h$ is one of the simplest, and responds well to local changes in density. Although it does not force the number of neighbours to remain constant my experience has shown that, in general, the number of neighbours does not vary very much. Miyama et al [41] show that using this method of updating $h$ in tests of slabs colliding head-on at high mach number can lead to particle clumping and overestimation of the density,
however this doesn’t present a problem in the simulations presented in this thesis. Benz [3] circumvented this problem by taking the time derivative of equation (2.15) to get

$$\frac{dh_i}{dt} = -h_i \nabla \cdot v_i$$  \hspace{1cm} (2.16)

which is added to the set of equations to be solved at each time step. These two methods for updating $h$ will be tested in section 4.1.

Alternate forms of updating $h$ generally involve keeping each particle’s number of neighbours constant. Nelson and Paploizou [54] achieve this by first building a list of near neighbours to particle $i$. They then find those particles on the list which are furthest from $i$, and take an average of the distances to $i$ to calculate $h$. This method is costly as it involves extra neighbour finding. Serna, Alimi and Chieze [70] predict a value of $h$, then count the number of neighbours within that radius. If the number found is not equal to the required number of neighbours, a new predicted value of $h$ is calculated. The procedure is repeated for each particle until the number of neighbours is right.

2.10 Neighbour searching

In order to calculate the rates of change of velocity and energy and in order to sum the densities, we must be able to calculate the interactions between neighbouring particles and, for that, we need to be able to determine which particles are neighbours. The simplest way of doing that is to begin with one particle then look at every other particle to determine how far away it is. A calculation with $N$ particles would require $N(N - 1)$ distance calculations. If the symmetry of the problem is taken into account, then this figure is halved. For very large computations, this is prohibitively slow.

Several other methods have been proposed to reduce the time it takes to search for neighbours. Two of the most common are link lists (see, for example, [20]) and the hierarchical tree method (see [17] and references therein). The tree method is particularly suited to SPH simulations with self-gravity – as the tree structure must be calculated for the gravitational calculations.
anyway. In this thesis, self-gravity is not considered, and one of the aims in developing the code was that it be simple to write, modify and add new physics to. With these factors in mind, the linked list method was chosen.

The link list method involves laying a grid over the computation region and assigning each particle to a grid cell. The particles are then linked in a link list structure, so that when we want to find neighbouring particles, we simply go to the appropriate grid cell number and check the link list to see which particles are nearby. It is impossible to determine an exact number of distance calculations involved here, since that depends strongly on the distribution of the particles. However, experimental evidence has shown this method to be far more efficient, particularly when the number of particles exceeds approximately one hundred.

As a simple example of how link listing works, consider an arrangement of particles in one dimension.

```
Cell Number  1  2  3  4  5  6
|   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   ●   

```

Particle 'a' is in the cell L where

\[
L = \text{Int} \left( \frac{x_a - x_{\text{min}}}{2h} \right) + 1
\]

We need two integer vectors: the head of chain which has length equal to the total number of cells (in this case, 6) and the link list which has length equal to the number of particles (in this case, 15). In order to construct the link list, we use the following algorithm:

Set both the link list and head of chain vectors to zero.
Loop over the particles, i:
  Calculate which cell, L, particle i is in
2.10. NEIGHBOUR SEARCHING

Set link list(i) = head of chain(L)
Set head of chain(L) = i

For the first few particles, the calculations are:

i=1
L=5
link(1)=hoc(5)=0
hoc(5)=1

i=2
L=4
link(2)=hoc(4)=0
hoc(4)=2

i=3
L=1
link(3)=hoc(1)=0
hoc(1)=3

i=4
L=5
link(4)=hoc(5)=1
hoc(5)=4

i=5
L=5
link(5)=hoc(5)=4
hoc(5)=5

This continues for the rest of the particles.

When we need to retrieve the particles to find neighbours, we look up the head of chain for the current cell number which gives us one of the particle numbers in that cell. We then follow the link list until we hit zero, which means there are no more particles in that cell. We repeat this for the adjacent
cells, and we have a list of particles which may be neighbours. In that list, there are only particles within a distance of $4h$ or less, so we don't need to search over more particles than necessary. When $h$ is not constant, we set the cell side length to be the maximum value of $h$. In this way, we ensure that we catch all possible neighbours, while still not searching over the entire simulation area. The grid is still a square or cubic grid, which means that the relevant code remains simple.

2.11 Summary

We have seen that SPH gives us a method of finding the values of a function and its derivatives on a set of disordered points, as well as a method of moving those points. We have also examined the equations we wish to model and translated them into their SPH forms. Two extensions to the SPH method – using a variable $h$ and a linked list for finding neighbours – which make the calculation of the SPH equations more efficient are incorporated into the code. They will be tested along with the other elements of the code in chapter 4. We will move on now to discussing the numerical methods which will allow us to step the fluid equations forward in time.
Chapter 3

Numerical Methods

The fluid equations are a set of coupled differential equations which we must solve. Chapter 2 introduced a method of calculating derivatives over a set of unordered interpolation nodes, or particles, and also examined ways of making those calculations more efficient. In this chapter we will use that information to numerically solve the fluid equations over a given time period.

The solution of a differential equation or set of equations, given a means of finding the derivatives, is something which has been well studied. This chapter begins with one of the more commonly used time stepping techniques in this field before progressing to an algorithm which is highly appropriate to equations of fluid flow in astrophysical settings and finally considers a method of maintaining accuracy and numerical stability throughout the simulation.

3.1 Time stepping

We now have a set of coupled differential equations, which we wish to solve with respect to time. There are a number of methods available for this class of problem, and one of the simpler is known as the Leapfrog predictor–corrector algorithm. This is based on Euler’s method, and goes as follows:
To step the equations:
\[
\frac{dv}{dt} = f \\
\frac{dr}{dt} = v
\]
from timestep $j$ to timestep $j+1$, where the size of the timestep is $\Delta t$

**PREDICT:**
\[
\hat{v}_{j+1} = v + \Delta t f_j \\
r_{j+1} = r_j + \Delta t v_j + \frac{1}{2} (\Delta t)^2 f_j
\]

**CORRECT:**
Calculate $f_{j+1}$ from $r_{j+1}$ and $\hat{v}_{j+1}$
\[
v_{j+1} = v_j + \frac{1}{2} \Delta t (f_{j+1} + f_j)
\]
\[
= \hat{v}_{j+1} + \frac{1}{2} \Delta t (f_{j+1} - f_j)
\]

Note that this is also a Runge-Kutta scheme.

While this method is adequate for moving the particles, it does not take into account the various time scales involved. For instance, in the acceleration equation, there are contributions to a particle’s acceleration from the pressure, viscosity and gravitation terms. The pressure and viscosity terms are naturally grouped, but the gravitation term is isolated, and the timesteps needed to resolve this part of the equation are about a tenth the size of those needed for the viscosity and pressure terms. It is possible to simply set the timestep to this lower value, however that is extremely inefficient, especially as the pressure and viscosity terms involve the slowest, most intensive computations in the program. The solution to this problem is to use the method of operator splitting proposed by Strang [72].
3.2 Strang Splitting

Strang splitting is used to aid the numerical integration of a differential equation where the differential operator is split into two or more parts. This scheme is appropriate for any such equation, but is applied most effectively in cases where the numerical difficulties are dominated by one particular term in the differential operator. That term may change on a different timescale to the rest, or it may be more difficult to solve, or it may introduce instabilities into the equations.

Consider the general differential equation

$$\frac{dx}{dt} = Cx \quad (3.1)$$

where $C$ may be a differential operator; a simple function; or a matrix of functions or differential operators. Now suppose that $C$ may be easily split into two component parts, $A$ and $B$, so that we can write

$$C = A + B, \quad (3.2)$$

and

$$\frac{dx}{dt} = (A + B)x. \quad (3.3)$$

We may now separate this equation into two subequations:

$$\frac{dx_A}{dt} = Ax \quad (3.4)$$

and

$$\frac{dx_B}{dt} = Bx. \quad (3.5)$$

Here, $x_A$ and $x_B$ are new variables: they correspond to what the solution of equation (3.1) would be if the only part of $C$ acting on the variable $x$ was (respectively) either $A$ or $B$.

In order to numerically integrate an equation such as (3.1) we could either directly solve the equation, or we could solve equations (3.4) and (3.5) alternately: Strang Splitting gives us the detail required to implement this
latter option. In this scheme, we begin by solving equation (3.4) over half a
timestep. The result of this is then fed into equation (3.5), which is solved
over a full timestep. Finally, the result of this is fed into equation (3.4), which
is again solved over half a timestep. In this way, we step from one timestep
to the next, getting the full solution to our general equation, equation (3.1),
at the end of each timestep.

As discussed at the end of this section and later in the thesis, there are
many advantages to using Strang Splitting. However, the scheme will not
be practical if it introduces inaccuracies or reduces the order of the original
numerical methods. In order to analyse these possibilities, we return to the
general differential equation:

\[
\frac{dx}{dt} = Cx.
\]

The distinction between the continuous function and its discrete approxima-
tions must be clear. In what follows, \( x = x(t) \) is the true value of the variable
\( x \) at time \( t \), while \( x^t \) is the approximation to \( x(t) \) obtained by stepping some
numerical scheme from the initial time to time \( t \).

We will also distinguish between the continuous differential operator and
its corresponding numerical counterpart. The numerical operator \( \widehat{C}_{\Delta t} \) is used
to step from \( x^t \) to \( x^{t+\Delta t} \), that is:

\[
x^{(t+\Delta t)} = \widehat{C}_{\Delta t}x^t
\]

The numerical operator in question may refer to any numerical integrator;
for example a Runge-Kutta scheme of any order, or the Leapfrog method.

We may now derive a relationship between the continuous and numerical
operators. If we assume that the numerical operator, \( \widehat{C}_{\Delta t} \) is accurate to
second order, we can use a Taylor expansion of \( x^{t+\Delta t} \):

\[
x^{(t+\Delta t)} \approx x^t + \Delta t \frac{dx^t}{dt} + \frac{(\Delta t)^2}{2} \frac{d^2x^t}{dt^2}
\]
and compare it with the numerical operator $\hat{C}_{\Delta t}$:

$$\hat{C}_{\Delta t} x^t = x^t + \Delta t \frac{dx^t}{dt} + \frac{\Delta t^2}{2} \frac{d^2 x^t}{dt^2}.$$  

However, the derivative operator corresponds to the continuous operator $C$, and so:

$$\hat{C}_{\Delta t} x^t = x^t + \Delta t C x^t + \frac{\Delta t^2}{2} C^2 x^t$$

$$\hat{C}_{\Delta t} = I + \Delta t C + \frac{\Delta t^2}{2} C^2.$$  \hspace{1cm} (3.6)

Note that this equation does not actually define the numerical operator, it merely provides a starting point for analysing its accuracy.

We will now return to our split equations (3.4) and (3.5), which were:

$$\frac{dx_A}{dt} = A x \quad \text{and} \quad \frac{dx_B}{dt} = B x.$$  

In the same way as $\hat{C}$ was defined, we define $\hat{A}$ and $\hat{B}$. We can also derive similar accuracy requirements, arriving at

$$\hat{A}_{\Delta t} = I + \Delta t A + \frac{\Delta t^2}{2} A^2$$  \hspace{1cm} (3.7)

$$\hat{B}_{\Delta t} = I + \Delta t B + \frac{\Delta t^2}{2} B^2.$$  \hspace{1cm} (3.8)

In the numerical method proposed by Strang [72], to get from one timestep to the next the operator $\hat{A}$ is applied with a half timestep, then the operator $\hat{B}$ with a full timestep, followed by $\hat{A}$ again with a half timestep. (Of course, the order of $\hat{A}$ and $\hat{B}$ is somewhat arbitrary, and using $ABA$ is generally equivalent to $BAB$.) That is,

$$x^{(t+\Delta t)} = \hat{A}_{(\frac{1}{2} \Delta t)} \hat{B}_{\Delta t} \hat{A}_{(\frac{1}{2} \Delta t)} x^t.$$  

In order to check that the accuracy of this scheme is at least the accuracy of the individual operators, we substitute in the relations (3.7) and (3.8),
which determine the accuracy of the operators \( \hat{A} \) and \( \hat{B} \).

\[
x^{(t+\Delta t)} = \hat{A}_{(\Delta t)} \hat{B}_{\Delta t} \hat{A}_{(\Delta t)} x^t
\]

\[
= \left( I + \frac{1}{2} \Delta t A + \frac{\Delta t^2}{8} A^2 \right) \left( I + \Delta t B + \frac{\Delta t^2}{2} B^2 \right) x^t
\]

\[
= \left( I + \Delta t (A + B) + \frac{\Delta t^2}{2} (A + B)^2 \right) x^t
\]

\[
= \left( I + \Delta t C + \frac{\Delta t^2}{2} C^2 \right) x^t.
\]

This is exactly the requirement given by equation (3.6), which shows that the Strang splitting scheme is at least as accurate as the accuracy of the numerical operators used to step each part of the equation.

As a concrete example, consider the fluid equation of acceleration,

\[
\frac{dv}{dt} = -\frac{1}{\rho} \nabla P + F_g + F_{\text{visc}}, \tag{3.9}
\]

in an accretion disc simulation. \( F_g \) are the forces due to gravity and \( F_{\text{visc}} \) are the viscous forces. We can estimate the relative time scales of those terms on the right hand side of (3.9) by evaluating \( \nu / \frac{dv}{dt} \) for each individual term. This gives us: the dynamical time scale

\[
t_{\text{dyn}} \sim \sqrt{\frac{R^3}{GM}};
\]

the radial sound crossing time scale

\[
t_s \sim \frac{R}{c_s};
\]

and the viscous time scale

\[
t_{\text{vis}} \sim \frac{R^2}{\nu}.
\]

For the alpha viscosity prescription, \( \nu = \alpha c_s H \), the viscous timescale be-
comes:

\[ t_{vis} \sim \frac{1}{\alpha} m^2 t_{dyn} \]

where \( m = v_k/c_s \) is the azimuthal mach number. In general \( t_{vis} > t_s > t_{dyn} \) and relative timescales may change throughout the disc.

These differences in timescale cause difficulties when numerically solving the equation. In order to resolve the dynamical timescale, the timesteps must be very small, yet that means taking many more steps than is necessary for the viscous timescale. Using Strang Splitting resolves these problems by stepping the equation in the following manner. To get from one timestep to the next, we first solve \( \frac{dv}{dt} = -\frac{1}{\rho} \nabla P \) with \( \frac{1}{2} \Delta t \), then \( \frac{dv}{dt} = F_g \), again with \( \frac{1}{2} \Delta t \), then \( \frac{dv}{dt} = F_{visc} \) with \( \Delta t \), then the pressure and gravity equations again, each with \( \frac{1}{2} \Delta t \). Of course, this order could be changed to whatever is practical, and the equation could be split into two parts (instead of three) by lumping together any of the operators.

When solving any of these equations, we can treat it as an isolated simulation. We have an equation, a set of initial conditions, and we must step from \( t \) to either \( t + \frac{1}{2} \Delta t \) or \( t + \Delta t \). This has a two significant implications. First, we may use any time stepping scheme we like, regardless of what is used to solve the other equations. We can choose the scheme and order of accuracy that is most appropriate to this equation, whether that be a predict-correct type scheme, a Runge-Kutta, or just something chosen at random from “Numerical Recipes” [62]. Second, we can use any timestep we like to solve the equation. We could use an appropriate physical step, such as a Courant condition, or just let an adaptive method choose its own step sizes.

3.3 Numerical Stability and Error Control

The overall numerical stability of the scheme will depend not only on the size of the timestep chosen, but on the behaviour of the system being simulated and on the stability of the numerical integrators used to step each part of the equation. Ideally, one would hope that the larger of the two natural timesteps was sufficiently small to ensure stability of the entire scheme,
however, this may not be guaranteed. Instead, a method of determining an appropriate timestep is required, preferably one which easily adapts to changing simulation conditions.

One simple solution is to generate some estimate of the error involved, then demand that this error be beneath a preset limit. In this way, not only is stability guaranteed, but an upper limit is also set on the overall error. We will describe here a method of generating an error estimate.

Starting at time $t$, the current value of the variable $x$ is $x^t$. Stepping to time $t + \Delta t$ gives $x^{t+\Delta t}$ which, according to the formalism in equation (3.2) is:

$$x^{(t+\Delta t)} = \hat{C}_{\Delta t} x^t$$

However, we have only demanded second order accuracy, so we can write that

$$x^{(t+\Delta t)} = Cx^t + O(\Delta t^3)$$  \hspace{1cm} (3.10)

where $C$ is the continuous operator. If, instead of taking one step of size $\Delta t$ we take two, each of size $\frac{1}{2} \Delta t$, we get:

$$x^{(t+\frac{1}{2}\Delta t+\frac{1}{2}\Delta t)} = Cx^t + O\left(\frac{\Delta t^3}{8}\right)$$  \hspace{1cm} (3.11)

Here, the current value of $x$ is written as $x^{(t+\frac{1}{2}\Delta t+\frac{1}{2}\Delta t)}$ in order to distinguish it from $x^{t+\Delta t}$.

We can now compare these two results to determine the error involved in moving from time $t$ to $t + \Delta t$:

$$x^{(t+\frac{1}{2}\Delta t+\frac{1}{2}\Delta t)} - \frac{x^{t+\Delta t}}{8} = Cx^t - \frac{C x^t}{8}$$

$$\therefore C x^t = \frac{8}{7} \left( x^{(t+\frac{1}{2}\Delta t+\frac{1}{2}\Delta t)} - \frac{x^{t+\Delta t}}{8} \right)$$  \hspace{1cm} (3.12)

Recalling that the error is the difference between using the continuous func-
tion $C$ to get to $t + \Delta t$ and using the numerical approximation:

\[
E_{t+\Delta t} = |\hat{C}x^t - Cx^t|
\]
\[
E_{t+\Delta t} = |x^{t+\Delta t} - Cx^t|
\]
\[
E_{t+\Delta t} = \left| x^{t+\Delta t} - \frac{8}{7} \left( x^{(t+\frac{1}{2}\Delta t + \frac{1}{2}\Delta t)} - \frac{x^{t+\Delta t}}{8} \right) \right|
\]
\[
= \frac{8}{7} \left| x^{t+\Delta t} - x^{(t+\frac{1}{2}\Delta t + \frac{1}{2}\Delta t)} \right|
\]

Similarly, the error involved in taking two half steps is:

\[
E_{t+\frac{1}{2}\Delta t + \frac{1}{2}\Delta t} = \frac{1}{7} \left| x^{t+\Delta t} - x^{(t+\frac{1}{2}\Delta t + \frac{1}{2}\Delta t)} \right|
\]

We now have an estimate of the local numerical error, as well as a method of determining whether the solution is convergent or divergent. As a bonus, we can also take the more accurate values at $t + \frac{1}{2}\Delta t + \frac{1}{2}\Delta t$. However, this method is costly. Generating the extra information increases the computation time by approximately one and a half. As a concession, we do not compute the error estimate every timestep. Instead, it is calculated after a certain number of timesteps. Every time an error estimate is generated, the number of steps to the next estimate is either increased or decreased, depending on how close the error is to the predefined level. Of course, if the error is ever larger than this, all steps since the last check must be abandoned, the timestep decreased, and error testing occurs every timestep until the levels are more acceptable.

\[1\] Although the value of the function $x$ computed at time $t + \frac{1}{2}\Delta t + \frac{1}{2}\Delta t$ is more accurate than that computed at time $t + \Delta t$, the combination of the two values given by equation 3.12 will not necessarily be even more accurate. This combination is an estimate based on simple Taylor expansions, and assumes that the $O(\Delta t^3)$ in equation 3.10 is equivalent to $8O(\Delta t^3)$ in equation 3.11, which is not necessarily the case. It is useful only for defining an upper bound to the error.
3.4 Summary

The timestepping technique and error control mechanism along with the method of finding a function's derivatives will allow us to solve the fluid equations presented in chapter 1. Before forging ahead to use these methods to solve interesting problems, we must first ensure that we will obtain the right answers. This is done with a series of comprehensive tests, which probe all aspects of the behaviour of the code and its underlying algorithms.
Chapter 4

Testing the Code

After any program is written, it must undergo rigorous testing. Initially, one must remove the obvious bugs in the code itself, and these are usually fairly simple tests. Next, one must ensure that the program is running roughly as expected. Finally, one should compare the output of various setups with either analytic results or previously published work, or both.

I will skip over the first two kinds of tests, as they are neither interesting nor particularly informative. Instead, I will focus on situations which have analytic solutions, and therefore verify the accuracy of both the SPH method, and of my code. I will then move on to addressing some concerns about the SPH method.

4.1 The Shock Tube test

The Riemann shock tube test is designed to test the ability of a hydrodynamic code to calculate viscosity and capture shocks. In this test, we simulate a long thin tube filled with gas. On one side the gas is at a high density and pressure and on the other side the gas is at a low density and pressure. There is a membrane separating the two regions which is removed at time zero. A shock then propagates into the low density region of the fluid while
a rarefaction wave propagates into the high density region. The fluid in the tube is considered to be adiabatic. Analytic solutions which follow the shock and rarefaction wave have been calculated in Harlow and Amsden [13].

This simulation begins with 9975 particles laid out on a square grid which covers from $x = -2$ to $x = 2$ along the length of the tube and $y = -0.125$ and $y = 0.125$ across the width of the tube. The spacing between the particles is equal in the $x$ and $y$ directions – it is 0.01. All particles begin with a smoothing length $h = 0.02$. Particles on the left hand side of the tube ($x < 0$) are assigned a mass of $1 \times 10^{-4}$, while those on the right are assigned a mass of $2.5 \times 10^{-5}$. These values give a density on the left of 1.0 and on the right of 0.25. The pressure on the left is set to be $P = 1.0$ and the on right is $P = 0.1795$.

The calculations begin from time 0, when the membrane is removed, and are run until after time 1.0. The viscosity is calculated using the prescription given by equation 2.13 with $\alpha = 1.0$ and $\beta = 2.0$. The boundaries of the tube are modelled using ghost particles: at each step, any particle within a distance of $2h$ of the edge of the tube is mirrored. The ghost particle lies outside the tube at the same distance to the edge as the mirrored particle, and has the mirrored particle’s values of density, pressure, and so on. Its velocity towards the edge across which it was mirrored is opposite the original particle’s.

There are three simulations performed here: the first has the smoothing length held constant throughout; the second has a variable $h$, which is updated via Benz’s equation (2.16); and the third also has a variable $h$, this time updated via equation (2.15).

The results shown in figures 4.1, 4.2 and 4.3 show the tube at time 0.5. The x-axis of each plot is the length of the tube. Simulation particles are represented by points on the plot, while the solid lines represent the analytical solutions. The top plot shows the velocity along the length of the tube, the second plot the density, the third the pressure, and the fourth the internal energy of the fluid in the tube. In each case, immediately after the shock there is some oscillation – this is most obvious in the velocity. The oscillations do not greatly effect the density, pressure or internal energy, though, all of
which are well represented by the particle values. The sharp edges of the shock front, contact discontinuity and rarefaction wave are all smoothed out somewhat as a result of the SPH smoothing kernel. This smoothing would decrease if particle number were increased.

In the first case, shown in figure 4.1, the post shock oscillation has the largest amplitude. Additionally, the agreement between the theoretical values and the simulated values of all physical quantities is not as good around the region of the rarefaction wave. The second case – using Benz’s equation to update $h$ – is shown in figure 4.2. The amplitude of the post shock oscillation is decreased, while the rarefaction wave is modelled more accurately. In this case, however, the post shock oscillations stretch further back into the shocked region, leading to some spread in the x-velocity around the contact discontinuity. The final case – that of updating $h$ with equation (2.15), is shown in figure 4.3. This simulation looks almost identical to the previous one, although the smearing around the contact discontinuity appears slightly smaller in magnitude.

4.2 The ring test

In this test, I will compare results from my code to analytic solutions obtained from thin disc theory. Most of this testing follows the work of Maddison, Murray and Monaghan [37] and Murray [52]. For a comprehensive coverage of the theory, see Pringle [63] - I will only state the necessary results here.

In this section, I want to test that both the movement of the particles and their density distribution behave as predicted. I will also demonstrate the need for a smoothing length that is able to adapt to changing conditions. To that end, the initial tests are run with a constant $h$ and subsequent tests are run with variable $h$. From thin disc theory, we find equations for the velocity and density of the gas at any given time and radius. Since this is a 2D case, we inspect the vertically averaged surface density, which is given
by:

$$\Sigma(r,t) = \frac{1}{6\nu t^{3/4}} \int_0^\infty \Sigma_0(r') r'^{5/4} \exp \left[-\frac{(r'^2 + r^2)}{12\nu t}\right] I_{1/4} \left(\frac{r'r}{6\nu t}\right) dr'$$  \hfill (4.1)

where $I_{1/4}(x)$ is a modified Bessel function of the first kind and $\Sigma_0(r)$ is the initial surface density as a function of radius. The value of $\nu$, the kinematic viscosity, is considered to be a constant. If we further assume that the angular speed of the gas is Keplerian, we can find the radial velocity of the gas as a function of the surface density:

$$v_r(r,t) = -3\nu \frac{\partial}{\partial r} \ln \left[r^{1/2} \Sigma(r,t)\right].$$  \hfill (4.2)

To complete the equations, we now need an initial surface density distribution. The simplest case would be a delta function distribution, however due to the smoothed nature of SPH it tends to smear out a delta function. It is better to model an initial function which is well represented by SPH. I have chosen to model a Gaussian function, given by:

$$\Sigma(0,t) = \begin{cases} \exp \left(-\frac{(r-r_e)^2}{l^2}\right) & r_{in} \leq r \leq r_{out} \\ 0 & \text{else} \end{cases}$$  \hfill (4.3)

where $r_{in}$ and $r_{out}$ are the inner and outer edges of the disc, $l$ is the Gaussian half-width and $r_e$ the location of the maximum density. This gives us equations for the surface density, and the initial radial velocity, which are as follows:

$$\Sigma(r,t) \approx \frac{1}{r^{3/4} \sqrt{\pi} 12\nu t} \int_{r_{in}}^{r_{out}} r'^3 \exp \left[-\frac{(r' - r_e)^2}{l^2}\right] \exp \left[-\frac{(r' r_e)^2}{12\nu t}\right] dr'$$  \hfill (4.4)

$$v_r(r,t) = -3\nu \frac{\partial}{\partial r} \ln \left(r^{3/2} \Sigma\right).$$  \hfill (4.5)

The equation (4.4) is an approximation of the true equation which is far simpler to calculate (see [37]). The approximation holds so long as $(r'r/6\nu t) \gg 1$. 
4.2. THE RING TEST

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>SPH viscosity coefficient</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Kinematic viscosity</td>
</tr>
<tr>
<td>$h$</td>
<td>Smoothing Length</td>
</tr>
<tr>
<td>$N$</td>
<td>Total number of particles</td>
</tr>
<tr>
<td></td>
<td>Number of rings</td>
</tr>
<tr>
<td>$r_{in}$</td>
<td>Inner edge of disc</td>
</tr>
<tr>
<td>$r_{out}$</td>
<td>Outer edge of disc</td>
</tr>
<tr>
<td>$r_c$</td>
<td>Centre of Gaussian</td>
</tr>
<tr>
<td>$l$</td>
<td>Gaussian half-width</td>
</tr>
</tbody>
</table>

Table 4.1: Summary of constants

The initial setup of the particles is a series of rings, each with approximately the same number density. In order to simulate the Gaussian distribution, I have given each particle a mass which corresponds to its density multiplied by the element of volume which it takes up. In figure 4.4 I have plotted the positions of a segment of the disc, the initial density distribution, and the radial and azimuthal velocity distributions. The values of important variables are given in table 4.1. The particles initially all have between 40 and 60 neighbours – those particles in the outer and inner rings have fewer neighbours. The first test is run holding $h$ constant for each particle.

Note here that the maximum density is slightly underestimated by SPH. This is an artifact of the finite size smoothing length. If the number of particles is increased and the smoothing length correspondingly decreased, the accuracy in this region is improved.

The results for the first test are shown in figure 4.5. This follows the evolution of the ring for 0.8 time units, a little more than one tenth of an orbit. The agreement here between SPH and thin disc theory (shown as solid lines) appears excellent, however this is not a long time to run such a simulation. Running to much later times, shown in figure 4.6, shows a substantial discrepancy between theory and SPH results, which appears to be growing.

I believe that this discrepancy can be explained as follows: as time increases, the particles are moving further apart. Their smoothing length,
however, remains constant. This means that particles influence each other less and less, and so the viscous contribution to their velocity drops off. Ultimately, each ring will be so far from its neighbour that only their different radial velocities determine how the band spreads. The rate of spreading drops below the level predicted as a result of viscosity, resulting in a similar discrepancy as shown in figure 4.6.

This problem may initially be solved by using more particles and placing the rings closer together. This delays the point where the influence of one ring on another drops off, allowing us to model the spreading band much further. However, at some point, the rings will spread far enough apart that the model breaks down.

Another solution is to allow the model to adaptively adjust to changing conditions. This is done by allowing the smoothing length to vary with time and space, as described in section 2.9. Results using this adjustment are shown in figures 4.7 and 4.8. The first of these shows the development at early times, and is much the same as the results from the test with constant $h$. The second shows the spreading of the ring at much later times, and the improvement is marked. Note that the setup is the same, and that the number of particles has not been increased. As in the previous section, using a variable $h$ has improved the accuracy and resolution of the simulation. This was also shown by Maddison et al [37] who investigated the instability of a single viscous ring. The use of variable $h$ is now common practise in most SPH simulations.

One final test is to check if angular momentum is conserved. This is done by calculating the angular momentum of each particle then averaging over the whole band. I have displayed the percentage change from the initial angular momentum for the tests with constant $h$ in figure 4.9 and with variable $h$ in figure 4.10. In order to check the longer term evolution, I have plotted the change in angular momentum of the variable $h$ case up to two orbits of the band. In both tests, the change in angular momentum is negligible (note the scale) and does not appear to show any trend in time.
4.3 A Rebuttal

In 2002 Imaeda and Inutsuka [22] published an article (hereafter referred to as II:2002) in which they claimed to have found serious errors in the density evolution of SPH simulations. They stated:

Evolutionary calculations of rotating gas flows around astrophysical objects with standard smoothed particle hydrodynamics (SPH) result in inaccurate evolutions of shear flows. The large density errors ($\Delta \rho \gg \rho$) emerge within one dynamical time of the system when we adopt the mean separation of the particles as the smoothing length, $h$.

To demonstrate this effect, they modelled several systems: a viscous ring with and without pressure effects and a cartesian shear box. I have modelled the first two of these tests with the code developed here, and have found no such errors.

The first test which II:2002 perform is to set up a small disc, similar to the ring test, except that the initial density corresponds to $\rho \propto r^{-1.5}$. They include pressure forces, and set $h$ to be a constant smoothing length equal to $3.7 \times 10^{-2}$. They mention that this is the mean inter particle spacing at $r = 1$. I have performed this test with my own code, and the results of this are displayed in figure 4.11. In this case, the simulation consisted of 18645 particles, and the viscosity parameters are both set to 0, following II:2002.

This figure is directly comparable to figure 1 in II:2002. As is obvious, this shows a very accurate evolution of the disc – certainly nothing like the errors which II:2002 reported.

In figure 4.12 I have replicated figure 2 of II:2002, showing the particle positions as time progresses. Again, this figure does not show any of the errors II:2002 reported. Furthermore, this figure also shows that the particles are definitely moving around the central object and that there is a radial shear, and SPH is still able to accurately model the evolution in density.

The second test of II:2002 is the same as the first except that in this test they have set the sound speed, and hence the pressure, equal to zero. In this
way, the only influence in the density calculation should be the Keplerian motion of the particles. Again, they find significant errors. Again, I have run the same simulation and found no errors. In fact, as could be expected, the results look almost identical to the first set. They are shown in figure 4.13.

The third test of II:2002 involved setting up a cartesian shear box. I have not done this test, as I believe that the first two are sufficient to demonstrate that the claims made by Imaeda and Inutsuka are not, in fact, due to an error in "the definition of density and the inaccurate solution of the continuity equation in standard SPH".

4.4 Summary

These tests have shown that SPH is well able to capture shocks and model viscous shearing. They also show that varying $h$ can greatly improve the resolution of some simulations, with little cost, and that the method chosen to update $h$ performs well.

As stated, the tests presented here are not the only ones performed – they are simply the most interesting. Further comprehensive tests have been performed on the neighbour finding algorithm, the timestepping and stability control algorithms, and on all other aspects of the method and code. The successful completion of these tests gives credibility to the results obtained in part II. From here on, simulations will be performed with the Strang splitting algorithm, the standard viscosity as given by equation (2.13), and with variable smoothing length, updated using equation (2.15).
Figure 4.1: The shock tube test: h constant
Figure 4.2: The shock tube test: $h$ variable, updated using equation (2.16)
Figure 4.3: The shock tube test: h variable, updated using equation (2.15)
Figure 4.4: The initial setup for the ring test

Figure 4.5: Results for the ring test
Figure 4.6: Results for the ring test, shown at times 1.0, 1.5 and 2.0.

Figure 4.7: Results for the ring test with variable h.
Figure 4.8: Results for the ring test with variable $h$, displayed at times 1.0, 1.5 and 2.0

Figure 4.9: Percentage change in angular momentum summed over all particles - constant $h$
Figure 4.10: Percentage change in angular momentum summed over all particles – variable h
Figure 4.11: The first simulation of Imaeda and Inutsuka: radial density profile. The y axis is density, the x axis is radius. Each particle in the simulation is represented by a cross.
Figure 4.12: The first simulation of Inada and Inutsuka: particle positions. The centre of the disc is at (0,0) and each particle is represented by a dot.
Figure 4.13: The second simulation of Imaeda and Imutsuka: radial density profile. The y axis is density, the x axis is radius. Each particle in the simulation is represented by a cross.
In part one, we examined the equations of fluid dynamics, and a method for solving these equations. First, we defined a way of interpolating a function from a set of disordered points. We also defined a way of obtaining the derivatives of that function. We then applied these methods to the fluid equations, arriving at a set of equations for the derivatives of density, velocity and pressure. An equation of state completed the set, giving us four equations in four variables.

Before considering ways of solving our fluid equations, we first turned to some simple ways of improving the efficiency of the derivative calculations. We allowed the kernel smoothing length to vary; thus allowing a simulation to react better to local changes in density by increasing or decreasing the resolution. We also examined a better way of determining which interpolation nodes (or particles) lie in the neighbourhood of any given point. This is necessary because each of the equations for calculating a derivative at a certain point requires knowledge of the surrounding points.

We then moved on to methods of stepping the fluid equations forward in time. First we looked at one of the simplest and most popular methods of solving initial value problems: the predictor-corrector method. While this would certainly work for many problems, it has its drawbacks; one of the biggest is its inability to efficiently handle equations with two or more implicit time scales. We turned then to the method of operator splitting proposed by Strang. We analysed its accuracy and proposed a method of choosing timestep size to ensure stability.

Finally, we tested the method and the code. We began with one of the simplest tests – the Riemann shock tube test. The code performed well, man-
aging to capture the shocks in the tube to a high degree of accuracy, with only minimal smoothing. Then we set up a ring of gas with a Gaussian density profile moving around a central object. The particles are given Keplerian velocity as well as a small radial velocity and allowed to interact. Viscosity spreads the ring out at a rate which may be determined analytically. Again, the code performed well, accurately following the evolution of the ring for quite some time. The final test was to repeat the tests of Imaeda and Inutsuka, who claimed to have found serious inaccuracies in SPH. These involved another ring, this time with a different initial density distribution which was in equilibrium and should therefore have remained the same. Again, the code performed well. The density, velocity and pressure profiles were accurately modelled for more than five orbital periods – twice as long as Imaeda and Inutsuka’s – and still there was no significant loss of accuracy.

And so we have a method of numerically solving the equations of hydrodynamics which has been proven to work. It accurately captures shocks, follows the viscous evolution of a ring and, when set in equilibrium, remains in equilibrium. We turn now, in part two, to the applications of this method. We will model discs around single stars and in binaries. We will model isothermal gases, then drop that assumption and model the temperature in the system. We will examine the inner boundary of discs, and how this effects the rest of the disc. In short – things are just beginning to get exciting.
Part II

Modelling Accretion Discs
Chapter 5

Discs Around Isolated Objects

The first, and probably simplest disc we turn to is a thin disc around an isolated object. Under a number of simplifying assumptions it is possible to derive a steady state solution for the gas flow in this kind of disc. In this chapter, I will outline the derivation of these solutions and model them with the SPH code. The derivation will follow the work of Shakura and Sunyaev [71], and Frank, King and Raine [10]. I will examine some of the problems surrounding the modelling of the inner boundary of the disc and propose a solution. Finally, the assumption that the gas in the disc acts isothermally will be dropped, and the full energy equation modelled. In this chapter, I will examine two forms of cooling – neither of which will be much more physical than an isothermal simulation. The point is to investigate the sensitivity of the energy equation, to examine the numerical problems associated with modelling a “full” energy equation and determine whether or not the Strang Splitting algorithm, developed previously, is robust enough to deal with these simulations.
5.1 Steady Thin Disc Theory

The central assumption of thin disc theory is, naturally enough, that the disc is (physically) thin – or, to be more precise, the scale height of the disc $H$ at a given radius $r$ is much less than that radius

$$H(r) \ll r.$$ 

Throughout this section, cylindrical polar coordinates will be used so that any point in the disc may be completely described by its radius $r$, azimuthal angle $\varphi$, and height $z$. The coordinates are oriented such that the orbital plane is the $z = 0$ plane. I will also assume that material in the disc orbits the central object at very close to the Keplerian speed with a small radial drift velocity. Under these assumptions it is possible to simplify the three dimensional system to a system of “two plus one” dimensions – that is, we can solve the equations in the orbital plane $(r, \varphi)$ and the equations perpendicular to that plane $(r, z)$ separately. This greatly simplifies the analysis.

Conservation Equations

The first thing to do is derive conservation equations for mass and momentum. Consider an annulus of disc material between $r_1$ and $r_1 + \Delta r$. It will contain total mass

$$2\pi r_1 \Delta r \Sigma$$

(5.1)

where $\Sigma$ is the surface density of the disc, given by integrating the mass density in the $z$-direction

$$\Sigma = \int_{-\infty}^{\infty} \rho(r, z) \, dz.$$ 

(5.2)

The same annulus of material will also contain an amount of angular momentum given by:

$$2\pi r_1 \Delta r \Sigma r_1^2 \Omega$$ 

(5.3)
5.1. STEADY THIN DISC THEORY

where $\Omega$ is the angular velocity of the material. Examining the net flow from neighbouring annuli will give the rate of change of mass and angular momentum in each annulus.

The rate of change of mass is given by the difference between the mass flow over the boundary at $r_1$ and the boundary at $r_2 = r_1 + \Delta r$, which, for each boundary, is equal to the velocity of the fluid in the radial direction, multiplied by the mass at that boundary:

$$\frac{\partial}{\partial t} \left( 2\pi r_1 \Delta r \Sigma \right) = v_r(r_1, t) 2\pi(r_1) \Sigma(r_1, t) - v_r(r_2, t) 2\pi(r_2) \Sigma(r_2, t) \quad (5.4)$$

Rearranging this equation gives:

$$2\pi r_1 \Delta r \left( \frac{\partial}{\partial t} (\Sigma) \right) = v_r(r_1, t) 2\pi(r_1) \Sigma(r_1, t) - v_r(r_2, t) 2\pi(r_2) \Sigma(r_2, t)$$

and so in the limit as $\Delta r$ tends to zero (as $r_1 \to r_2$),

$$r_1 \frac{\partial}{\partial t} (\Sigma) + r \frac{\partial}{\partial r} (r \Sigma v_r) = 0 \quad (5.5)$$

The rate of change of angular momentum is derived in a similar fashion, except there is an additional term due to the effect of one annulus exerting a viscous torque on another. At this stage, I will simply label the torque which the ring at $r_1$ exerts on the ring just inside it as $G(r_1, t)$. The angular momentum conservation equation is:

$$r \frac{\partial}{\partial t} (\Sigma r^2 \Omega) + \frac{\partial}{\partial r} (r \Sigma v_r r^2 \Omega) = \frac{1}{2\pi} \frac{\partial G(r)}{\partial r}. \quad (5.6)$$

**Viscosity and the $\alpha$-prescription**

The viscous torque above is one of the mechanisms by which the disc transports angular momentum, but how can it be quantified? Unfortunately, very little is known about viscosity, particularly in astrophysical situations. We do
know, though, that viscosity may be generally due to one of four things: the existence of a magnetic field; turbulence; molecular viscosity; and radiative viscosity. Of these, it is only the first two which will substantially contribute to the viscosity in an accretion disc – both molecular and radiative viscosity will be negligibly small.

For the case of magnetic viscosity, the shear will be determined by the magnetic energy in the disc. If the magnetic energy is greater than the thermal energy of the matter, then the magnetic stresses and efficiency of angular momentum transport will be high enough to ensure that the accretion happens radially – this is clearly not the case in an accretion disc, so we will assume that the energy of the magnetic field does not exceed the thermal energy:

$$\frac{\mathcal{H}^2}{8\pi} < \frac{c_s^2}{2}$$

(c_s is, of course, the sound speed). This gives magnetic stresses of:

$$w_{mc} \sim -\rho c_s^2 \left( \frac{\mathcal{H}^2}{4\pi \rho c_s^2} \right).$$  \hspace{2cm} (5.8)

In the case of turbulent viscosity, in order to determine the stress tangential to the flow, we must consider the speed and maximum size of turbulent motions. The speed, \(v_t\), must be less than the sound speed in the disc – turbulent motion at speeds higher than the sound speed will cause shocks to develop, which will in turn heat the gas, and the turbulent speed will become smaller than the sound speed. The length scale of turbulence may be anywhere up to the scale height of the disc, \(H\). On large scales, the turbulence will be homogeneous and isotropic, and so the viscous stresses may be approximated by the formula which gives viscous stress for molecular viscosity. The stress is then given by the viscosity multiplied by the velocity gradient, and we may substitute the turbulent viscosity for the molecular one, ending with an order of magnitude estimate of

$$w_{re} \sim -\rho c_s^2 \frac{v_t}{c_s}.$$  \hspace{2cm} (5.9)
5.1. STEADY THIN DISC THEORY

Combining these two equations, (5.8) and (5.9) gives

\[-w_{r\varphi} \sim \rho c_s^2 \left( \frac{v_t}{c_s} + \frac{\mathcal{H}^2}{4\pi \rho c_s^2} \right) = \alpha \rho c_s^2 \]  

(5.10)

It is clear that, in all reasonable accretion disc situations, \( \alpha \) must be less than one – usually, it is likely to be much less.

We are now in a position to give an equation for the torque, \( G(r) \). Torque is of course given by the moment of the force per unit area, multiplied by the area over which the force is applied and by \( r \) to get the total torque, so for a ring at radius \( r \), we get that

\[ G = 2\pi r^2 w_{r\varphi} H \]
\[ = 2\pi r^2 \alpha \rho c_s^2 H \]
\[ = 2\pi r^2 \alpha \Sigma c_s^2. \]

(5.11)

Note that here I have used the relation \( \Sigma = \rho H \). While this is not true for a general \( \rho(z) \), it will be for a vertically averaged \( \rho \), and since we are using vertically averaged variables throughout this derivation, a vertically averaged density will suffice.

**Steady Discs**

In general, the rate at which material spirals in through the disc will be quite slow, and external conditions could well change much more quickly. To see this, consider the mass conservation equation, (5.5) – a rough estimate of the time scale under which the surface density will change is

\[ t \sim \frac{r}{v_r}. \]

So long as the radial velocity is small, this timescale will be large, and so long as the timescale is large, a steady state may be reached in the disc.

Equations for a steady state disc may be obtained from the two conservation equations, by setting the time derivatives to zero. From the mass
conservation equation, we get
\[ \frac{\partial}{\partial r} r \Sigma v_r = 0 \tag{5.12} \]
which is simply a requirement that the mass inflow through the disc be constant. We can define the mass transfer rate, or the amount of mass flowing through the disc as:
\[ \dot{M} = -2\pi r \Sigma v_r \tag{5.13} \]
where \( v_r \) is defined to be the outward radial velocity, so \( \dot{M} \) is positive for radial inflow.

From the angular momentum conservation equation;
\[ (r \Sigma v_r r^2 \Omega) = \frac{G}{2\pi} + \frac{C}{2\pi} \tag{5.14} \]
the constant \( C \) comes from the integration. \( C \) may be determined by assuming that at some point, \( r_0 \), the viscous torque disappears. Consider the object at the centre of the disc: in the cases in which we are interested it will be a white dwarf, though the following reasoning also holds true for many other cases. The white dwarf will be rotating with some angular velocity, \( \Omega_* \) at the equator. This angular velocity must be less than the Keplerian – otherwise the star would break up. This means that the speed of the material in the disc steadily increases, according to the Kepler Law, then quickly drops to \( \Omega_* \) within a small boundary layer surrounding the star. Of course that implies that there is a point close to the star at which \( \Omega' \) is equal to zero, and the torque must be zero, or close to it. Therefore,
\[ C = 2\pi \left( r_0 \Sigma v_r r_0^2 \Omega(r_0) \right) \]
\[ C = \dot{M} \Omega(r_0) r_0^2. \tag{5.15} \]
Substituting back into equation (5.14),

$$2\pi \left( r \Sigma v_r r^2 \Omega \right) = 2\pi r^2 \alpha \Sigma c_s^2 + \dot{M} \Omega (r_0) r_0^2$$

and using the Keplerian angular velocity at \( r_0 \),

$$2\pi \alpha \Sigma c_s^2 = \dot{M} \Omega \left( 1 - \left( \frac{r_0}{r} \right)^{\frac{3}{2}} \right). \quad (5.16)$$

The rotation of the disc supports the disc material against the gravitational force of the central object, but only in the radial direction. In the vertical direction, it is only the pressure forces which support the gas against the vertical component of gravity. Thus, the height of the disc is determined by the equation of hydrostatic equilibrium:

$$\frac{1}{\rho} \frac{\partial P}{\partial z} = \frac{\partial \Phi}{\partial z}. \quad (5.17)$$

Using the original assumption that the disc is thin, the typical scale height of the disc is given by

$$H = \frac{c_s}{v_\varphi} R = \frac{c_s}{\Omega}$$

The pressure in the disc is given by the normal equation for an ideal gas:

$$P = \frac{\rho k T}{\mu m_p} + \frac{4\sigma}{3c} T^4$$

which also gives an equation for the sound speed

$$c_s^2 = \frac{P}{\rho}.$$

Finally, we need an energy equation to complete the set. Energy is generated in the disc by viscous dissipation and liberated by an energy flux in the vertical direction. Since the disc is in a steady state, these two effects
must balance each other. The viscous dissipation is simply given by

\[ Q = \frac{G \Omega}{4\pi R} \frac{d\Omega}{dr} = \frac{1}{2} R \alpha c_s^2 \Sigma \frac{d\Omega}{dr} = \frac{3}{8\pi} \dot{M} \frac{GM}{r^3} \left( 1 - \left( \frac{r_0}{r} \right)^{\frac{3}{2}} \right) \] (5.18)

while the energy liberated is related to the opacity of the disc and the black-body energy equation – we are assuming that the disc is optically thick –

\[ Q = \frac{4 \sigma T_c^4}{3\tau} \]

where \( T_c \) is the central temperature.

Collecting all of these equations together, along with an equation for the optical depth, gives us a closed set of algebraic equations.

1. \[ \frac{4 \sigma T_c^4}{3\tau} = \frac{3}{8\pi} \dot{M} \frac{GM}{r^3} \left( 1 - \left( \frac{r_0}{r} \right)^{\frac{3}{2}} \right) \]
2. \[ c_s^2 = \frac{P}{\rho} \]
3. \[ P = \frac{\rho k T}{\mu m_p} + \frac{4\sigma}{3c} T^4 \]
4. \[ h = \frac{c_s}{v_\phi} \]
5. \[ 2\pi \alpha \Sigma c_s^2 = \dot{M} \Omega \left( 1 - \left( \frac{r_0}{r} \right)^{\frac{3}{2}} \right) \]
6. \[ \rho = \frac{\Sigma}{h} \]
7. \[ \tau = \tau(\Sigma, \rho, T_c) \]

Given an expression for the optical thickness of the disc, the solution of this set of equations is straightforward. Shakura and Sunyaev divide the disc into
5.1. STEADY THIN DISC THEORY

three regions:

a) radiation pressure is dominant and electron scattering on free electrons gives the main contribution to the opacity.

b) gas pressure is dominant; electron scattering is still the main contribution to the opacity.

c) gas pressure is again dominant, but free-free absorption and other mechanisms determine the opacity.

These three regions correspond respectively to the inner, middle and outer regions of a general accretion disc. For the purposes of this thesis, the inner region may be ignored – the radius where radiation pressure begins to dominate over gas pressure is smaller than the radius of a white dwarf.

For the outer part of the disc, where the opacity is dominated by free-free absorption, we can set the optical depth to be

\[
\tau_{ff} = \Sigma \kappa_{n,ff} = 5 \times 10^{24} \Sigma \rho T^{-7/2}.
\]  

(5.19)

where \( \kappa_{n,ff} \) is the Kramer’s Law approximation to the Rosseland mean opacity.

Using dimensionless variables

- \( m = \frac{M}{M_\odot} \)
- \( \dot{m} = \frac{\dot{M}}{10^{16} \text{gs}^{-1}} \)
- \( r = \frac{r}{10^{10} \text{cm}} \)

and dropping the radiative term from the equation for pressure, the system is algebraic and straightforward to solve. If there isn’t a decent symbolic algebra package at hand, combine equations 1, 6 and 7 to get

\[
\frac{4}{3} \sigma T^{15/2} (5 \times 10^{24} \Sigma^2 H^{-1}) = \frac{3}{8} \frac{G \dot{m} \dot{n}}{\pi r^3} f^{1/4}
\]
where \( f^4 = 1 - \left( \frac{r_*}{r} \right)^{1/2} \). Then use equations 2, 3 and 4 to eliminate \( h \), and equation 5 to eliminate \( \Sigma \). Without the numerical factors this leaves:

\[
T^8 \propto \Sigma^2 \dot{m} m^{3/2} r^{-9/2} f^4
\]

and

\[
\Sigma \propto \alpha^{-4/5} \dot{m}^{7/10} m^{1/4} r^{-3/4} f^{14/15}.
\]

The others all follow fairly easily, and the full Shakura-Sunyaev disc solution for the outer parts of the disc is:

\[
\Sigma = 5.2 \alpha^{-4/5} \dot{m}^{7/10} m^{1/4} r^{-3/4} f^{14/5} g \text{ cm}^{-2} \quad (5.20)
\]

\[
H = 1.7 \times 10^8 \alpha^{-1/10} \dot{m}^{3/20} m^{-3/8} r^{9/8} f^{3/5} \text{ cm} \quad (5.21)
\]

\[
\rho = 3.1 \times 10^{-8} \alpha^{-7/10} \dot{m}^{11/20} m^{5/8} r^{-15/8} f^{11/5} g \text{ cm}^{-3} \quad (5.22)
\]

\[
T = 1.4 \times 10^4 \alpha^{-1/3} \dot{m}^{3/10} m^{1/4} r^{-3/4} f^{6/5} \text{ K} \quad (5.23)
\]

\[
\tau = 190 \alpha^{-4/5} \dot{m}^{1/5} f^{4/5} \quad (5.24)
\]

\[
v_r = 2.7 \times 10^4 \alpha^{4/5} \dot{m}^{9/10} m^{-1/4} r^{-1/4} f^{-14/5} \text{ cm s}^{-1} \quad (5.25)
\]

with \( f = \left( 1 - \left( \frac{r_*}{r} \right)^{1/2} \right)^{1/4} \).

The variables here are scaled to be appropriate to discs around white dwarfs, so it is simple to get order of magnitude checks on some of the assumptions made in deriving these solutions. The first assumption was that \( H(r) \ll r \). Since \( H \sim 10^8 \), and \( r \sim 10^{10} \), clearly \( H/r \ll 1 \). The optical depth is greater than one, so the assumption that the disc is optically thick holds. The ratio of pressure due to radiation and gas pressure will be

\[
\frac{P_{\text{rad}}}{P_{\text{gas}}} \sim 2.8 \times 10^{-3}
\]

so the assumption that gas pressure dominates also holds.

In the middle regions of the disc, where the opacity is given by electron scattering, the optical depth is given by

\[
\tau_{es} = \Sigma 0.4 \text{ cm}^2 \text{ g}^{-1}.
\]
The solutions are obtained in the same manner as above. Since the opacity due to free-free absorption, $\kappa_{R,ff}$, is given by:

$$\kappa_{R,ff} = 36\dot{m}^{-1/2}m^{1/4}r_{10}^{3/4} f^{-2}$$

which dominates whenever

$$\frac{\kappa_{R,ff}}{\kappa_{es}} = 36\dot{m}^{-1/2}m^{-1/4}r_{10}^{3/4} f^{-2} > 0.4$$
$$r_{10}^{3/4} > 0.01\dot{m}^{1/2}m^{-1/4} f^2$$
$$r > 2.5 \times 10^7 \dot{m}^{2/3}m^{-1/3} f^{8/3} \text{ cm.}$$

For all reasonable values of $\dot{m}$, this radius is greater than the radius of a white dwarf, so we may safely ignore the solutions to the "middle" section of the disc.

## 5.2 The Inner Boundary

Between the star and that point in the disc when the angular velocity of the disc material departs from the Keplerian value there exists a boundary layer. This layer is thought to be very thin and the normal disc equations no longer hold there. For example, as $r \rightarrow r_0$, from equation (5.16) $\Sigma \rightarrow 0$, and $v_r \rightarrow \infty$; this is clearly not physical.

In this thesis, I am only modelling the disc itself so may blithely ignore the boundary layer. However, this still leaves some computational problems. At what point does disc modelling stop and the boundary layer begin? And what happens to the disc particles as they reach this radius?

The first question is answered somewhat simplistically – I assume that the boundary layer exists for $r \leq 1.1r_*$, and ensure that I only model the disc outside this radius. Often, the radius of the inner boundary of the disc is set to be much larger due to computational concerns. One cannot model a disc with an inner radius greater than twice the smoothing length, for example, otherwise particles on opposite sides of the central object may be able to "see" each other.
The second question is much more difficult to answer. The simplest thing to do is to remove the particles once they drop over a predefined inner radius, and this is the most commonly applied solution. Using this method, the structure of the disc suffers effects up to a distance of approximately 2R. This effect is demonstrated in figure 5.1, which shows the inner region of a disc simulation about an isolated one solar mass object. The radius, density and time are all in scaled units – radius is scaled such that the disc normally extends out to a radius of 1, and time so that a particle at radius 1 on a Keplerian orbit will have a period of 2π. The crosses in this plot are the results of the simulations, while the solid line is the steady thin disc solution.

![Figure 5.1: Normal inner boundary condition. x axis is radius, y axis is density.](image)

Normally, this effect may not be troubling. Any simulations would be worthless in the very inner regions of the disc, but you could just cut them out when drawing up the results. Unfortunately, this is not the case when the disc being modelled no longer follows a simple cooling law – as is the case later in this chapter. If the heating or cooling term, or the balance between them, is particularly sensitive to the density or pressure, then this inner boundary problem may no longer be so trivial. Later in this chapter we will see that the heating term, which is mainly due to the viscosity, depends on the gradient of the velocity – which for a Keplerian disc increases for decreasing r. The cooling depends directly on some power of density, either positive or negative. This means that the inner regions of the disc heat up as they should: the effect of the inner boundary does not alter the velocity gradient.
Unfortunately, as the density is decreased below its normal level, the cooling also becomes either less or more efficient, allowing the particles in the inner regions of the disc to heat to high temperatures, or to cool to unphysically low temperatures. This introduces numerical instabilities making it almost impossible to continue to model the disc. One can no longer simply throw out the results from the inner region.

In this thesis, I have developed a new method of modelling the inner boundary which resolves some of these problems. I define an inner boundary \(r_{buffer}\) such that \(2h < (r_{buffer} - r_{in})\). When a disc particle moves inside this boundary, its pressure and density are held static. It continues to move under the influence of viscosity and gravity, and continues to interact with other particles, both inside and outside \(r_{buffer}\). Such a particle is termed a buffer particle. Once a buffer particle moves over the radius \(r_{in}\), it is deemed to have been accreted, and is removed from the simulation. This layer between \(r_{in}\) and \(r_{buffer}\) provides a buffer region so that the disc can no longer “feel” the zero pressure hole at the centre.

Figure 5.2: Buffer particle inner boundary condition. x axis is radius, y axis is density.

Figure 5.2 shows the results of a simulation run with this buffer particle method at the inner boundary. All other parameters are identical to the simulation shown in figure 5.1. Again, the solid line is the Shakura–Sunyaev solution. Buffer particles are not plotted – they lie just inside the radius of the inner result set and have approximately the same density.
5.3 The Energy Equation

When an isothermal disc is modelled, it is implicitly assumed that any heating in the gas will be locally and immediately lost somehow. This makes numerical modelling much easier, since the energy equation need not be modelled at all. The problem with the energy equation is that it can be very sensitive to the balance between heating and cooling, and can tend to render the entire model numerically unstable. Unfortunately, while isothermal simulations are interesting from the point of view of the dynamics of discs, it is not particularly physical to ignore thermodynamics entirely, and the temperature structure can play a large part in the dynamics of the disc.

If all of the physics were taken into account, the cooling term would be quite complex – ideally one might want to model the full radiative transfer equation in three dimensions, taking into account local disc properties, as well as illumination from the primary star and boundary layer. Unfortunately, this goal remains out of reach – even if one could successfully develop an appropriate algorithm, the computing time would be prohibitive. Some assumptions and generalisations must be made.

Recently, several attempts have been made to include the effects of radiative transfer in SPH simulations. Both Whitehouse and Bate [77] and Bastien, Cha and Viau [2] have included radiation using the flux limited diffusion approximation. Oxley and Woolson [57] have devised a scheme whereby the fluid equations and radiative transfer equations are solved individually; radiative transfer is calculated between hydrodynamic timesteps and the resulting temperatures are applied to the SPH particles. This method allows easy incorporation of radiating sources, however it is extremely computationally expensive, and assumes that matter and radiation temperatures are equivalent.

The SPH form of the energy equation was introduced in section 2.3.3 and the viscous heating term in section 2.7. Together, they are

\[
\frac{dU}{dt} = \frac{1}{2} \sum_b m_b \left( \frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2} + \Pi_{ab} \right) v_{ab} \cdot \nabla_a W_{ab}.
\]
5.3. THE ENERGY EQUATION

This is, of course, the SPH equivalent to the fluid equation:

\[
\frac{dU}{dt} = -\frac{P}{\rho} \nabla \cdot \mathbf{v} + \frac{dU}{dt_{visc}}
\]  

(5.26)

and there is a term missing – the cooling term. The equation will be:

\[
\frac{dU}{dt} = -\frac{P}{\rho} \nabla \cdot \mathbf{v} + \frac{dU}{dt_{visc}} - \frac{dU}{dt_{cool}}
\]  

(5.27)

with some appropriate choice for \( \frac{dU}{dt_{cool}} \).

Heat will be lost from the disc in the form of radiation and so probably the most important factor in determining the cooling term is the optical depth, \( \tau \). If \( \tau \ll 1 \), then any photons emitted by the gas are released immediately and not absorbed. In this case, the rate of energy lost to the disc per unit volume will be solely due the the emissivity, \( j \),

\[
\frac{du}{dt_{cool}} = 4\pi j \text{ erg s}^{-1} \text{ cm}^{-3}.
\]

If we assume the disc is nearly fully ionised, then a good approximation for the emissivity is that given by thermal Bremsstrahlung:

\[
4\pi j = 6.8 \times 10^{-38} \int_{0}^{\infty} n^2 T^{-1/2} \exp \left( \frac{-h\nu}{k_B T} \right) g_{ff} d\nu \text{ erg cm}^{-3} \text{ s}^{-1} \text{ Hz}^{-1}
\]

\[
\simeq 1.4 \times 10^{-27} n^2 T^{1/2} \bar{g}_B \text{ erg s}^{-1} \text{ cm}^{-3}
\]

where \( n \) is the number density of ions and electrons, \( g_{ff} \) is the velocity averaged Gaunt factor, and \( \bar{g}_B \) is the frequency average of \( g_{ff} \). (See, for example [59] or [66].) Of course, the number density is the mass density divided by
the mass, and so

\[ \frac{du}{dt_{\text{cool}}} = 1.4 \times 10^{-27} \left( \frac{\rho}{m} \right)^2 \frac{T^{1/2}}{\bar{g}} \text{ erg s}^{-3} \text{ cm}^{-3} \]

\[ = 1.4 \times 10^{-27} \left( \frac{1}{m} \right)^2 \rho \frac{T^{1/2}}{\bar{g}} \text{ erg s}^{-1} \text{ g}^{-1} \]

The frequency averaged Gaunt factor is a function of \( T \), however it only varies between about 1.1 and 1.5 – 1.2 is usually a good representative value. Substituting equation 2.6 for the temperature, we have

\[ \frac{dU}{dt} = 5.52 \times 10^{16} \rho U^{1/2} \text{ erg s}^{-1} \text{ g}^{-1} \]

which may then be scaled to one solar mass and 10\(^10\) cm to give the scaled version

\[ \frac{dU}{dt} = 8.27 \times 10^{13} \rho U^{1/2}. \] \quad (5.28)

On the other hand, if \( \tau \gg 1 \) then the energy lost is the same as that obtained in section 5.1:

\[ \frac{du}{dt_{\text{cool}}} = \frac{4\sigma T^4}{3\lambda} \text{ erg s}^{-1} \text{ cm}^{-3} \]

\[ \frac{dU}{dt_{\text{cool}}} = \frac{4\sigma T^4}{3\kappa \Sigma \rho} \text{ erg s}^{-1} \text{ g}^{-1}. \]

The opacity, \( \kappa \) was given by equation (5.19) as

\[ \kappa = 5 \times 10^{24} \rho T^{-7/2} \text{ cm}^2 \text{ g}^{-1} \]

and so we have that in the limit of very large optical depth

\[ \frac{dU}{dt_{\text{cool}}} = 1.512 \times 10^{-29} \rho^{-2} T^{15/2} \Sigma^{-1} \text{ erg s}^{-1} \text{ g}^{-1}. \]

Again substituting equation 2.6 for the temperature, and scaling we have

\[ \frac{dU}{dt_{\text{cool}}} = 4.386 \times 10^{-2} \rho^{-2} U^{15/2} \Sigma^{-1}. \] \quad (5.29)
The relationship between the density $\rho$ and the surface density $\Sigma$ is given by thin disc theory as

$$
\Sigma = \rho H
= \rho \frac{c_s}{\Omega}
$$

Clearly, these two different cooling terms have very different dependence on the temperature and density in the disc, and so we could expect them to behave differently. The optically thin cooling law is directly proportional to the density, so the higher density inner regions of the disc ought to cool more. These are the regions which also heat more quickly, though, and so the balance is difficult to pre-determine. The optically thick cooling law is inversely proportional to the density cubed, so here, one might expect the inner regions of the disc to heat quite rapidly. However, the cooling is also directly proportional to $U^{15/2}$ and so the more the disc heats, the more it cools – perhaps the two effects will balance each other.

5.4 Disc Simulations

In order to examine the effects of these two different cooling laws, I have run a number of simulations of accretion discs around an isolated one solar mass object, using the code developed throughout part I. This code utilises the Strang Splitting algorithm and error control techniques described in chapter 3. Each simulation begins in the Shakura-Sunyaev steady disc state, with thermal energy, surface density, radial velocity, and so on, determined by equations (5.20) to (5.25). All simulations are performed in two dimensions. I have also defined a parameter, $\lambda$, which determines how efficiently the disc is able to cool. A $\lambda$ value of 1 means that the disc cools exactly as described in the cooling law. $\lambda < 1$ means that the disc cools less, and $\lambda > 1$ that the disc cools more. The cooling term implemented is then

$$
\frac{dU_{SPH}}{dt} = \lambda \frac{dU}{dt}
$$
where \( \frac{dU}{dt} \) is given by 5.29 for the optically thick simulations and 5.28 for the optically thin simulations.

There are two sets of simulations – one for each of the cooling laws. Within each set there are several simulations with different cooling efficiencies, \( \lambda \). The names of each simulation, and the values of \( \lambda \) are listed in table 5.1.

Each simulation has the same initial setup. There are 25 concentric rings of particles, starting at \( r = 0.084 \) and ending at \( r = 0.4 \). The width between the rings is approximately 1.1h. In each ring, the number of particles varies so that the interparticle spacing is roughly the same as the width between the rings. The Shakura-Sunyaev density is calculated for each ring, then the mass of the particle set to be the density at that radius, multiplied by the radius and the square of the interparticle spacing. In this way, when the density is calculated by summing over the particles, it returns the correct Shakura-Sunyaev value. The inner boundary is computed using the buffer particle method described above, and the smoothing length is allowed to vary. Each simulation has 21,998 particles, with each particle initially having a smoothing length of \( h = 0.011 \). This gives each particle approximately 80 neighbours. The number of neighbours doesn’t change appreciably throughout each run, as the density profile doesn’t change.

The thermal timescale of a disc is the time taken for the disc to re-adjust to a change in the energy dissipation or viscous heating rate. It can be defined as the heat content per unit area of the disc divided by the dissipation rate per unit area. The heat content is \( \approx \Sigma c_s^2 \) while the viscous dissipation is

<table>
<thead>
<tr>
<th>Simulation</th>
<th>Cooling Law</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thick1</td>
<td>Thick</td>
<td>0.1</td>
</tr>
<tr>
<td>Thick2</td>
<td>Thick</td>
<td>1</td>
</tr>
<tr>
<td>Thick3</td>
<td>Thick</td>
<td>10</td>
</tr>
<tr>
<td>Thin1</td>
<td>Thin</td>
<td>0.5</td>
</tr>
<tr>
<td>Thin2</td>
<td>Thin</td>
<td>1.0</td>
</tr>
<tr>
<td>Thin3</td>
<td>Thin</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 5.1: Summary of constants
given by equation (5.18). This gives

\[ t_{th} \sim \frac{1}{v_\phi \alpha}. \]

For a simulation where the azimuthal velocity varies between around 0.2π and 2π, this gives a thermal timescale around 0.16 – 1.6 with \( \alpha = 1 \). With this in mind, each simulation was run for at least two time units, by which time the temperature structure should have settled down to close to an equilibrium state. Of course, if the change in energy loss renders the disc physically unstable, it will never settle into an equilibrium state and the simulation quickly becomes numerically unstable.

For each simulation, I have displayed the disc’s thermal energy and density for a number of time frames. In order to do this, the physical quantity is first interpolated from the particles onto a grid. The grid is then plotted in the \( xy \) plane with each location coloured according to the value of either the density or thermal energy at that position. The coloured plots have the advantage that any azimuthal variations in the disc are clearly shown, however it is difficult to show the radial structure in a quantitative way in this kind of plot. Therefore, underneath each disc, I have also plotted the value of the thermal energy or density of each particle as a function of radius. Each particle is plotted as a cross, and the initial values as a solid line.

Note that due to the interpolation procedure, the scales of the two kinds of plots differ. The drop in density and temperature at the centre of the coloured plots is not due to the inner boundary in the simulations, but to the inner boundary in the interpolation procedure. The slight diffraction patterns are also a result of the interpolation onto the grid.

**Optically Thick Discs**

The first set of simulations were run with the optically thick cooling law. One might expect that the temperature and density structure would not change greatly from the Shakura-Sunyaev values, as those equations were calculated assuming that the disc was optically thick, and indeed, this is the case.

The first of the simulations was run with under-efficient cooling, so that
the disc didn’t cool as well as it should. The cooling efficiency parameter, \( \lambda \) was 0.1, so cooling was 10\% efficient. The thermal energy and density are plotted in 5.4 and 5.5 respectively. This resulted in a substantially higher temperature in the inner regions of the disc, but didn’t really effect the outer regions. As expected, the density remained unchanged.

Running with the proper cooling efficiency still heats the inner regions of the disc slightly, but at the same time, cools the outer regions so that, overall, the disc remains at roughly the same temperature. Again, as expected, the density is unchanged. This simulation is plotted in figures 5.6 and 5.7.

The third simulation, shown in figures 5.8 and 5.9, modelled a disc with over-efficient cooling (10 times the normal cooling efficiency). This disc behaves in much the same way as the previous, with the inner regions heating slightly and the outer regions cooling.

When the disc is optically thick, it is remarkably insensitive to the efficiency of the cooling. I have tested cooling efficiency factors of between \( 10^{-2} \) and \( 10^{2} \), and the resulting structure varies only a little. In general, the density doesn’t change, and the shape of the temperature curve remains the same, while the mean temperature increases or decreases slightly.

**Optically Thin Discs**

The second set of simulations were run with the optically thin cooling law. In this case, we could expect some change in the temperature structure, as we have broken one of the key assumptions of Shakura–Sunyaev theory. The density structure, however, should not change. This is because the viscous timescale, which governs how long it takes for density changes to propagate through the disc, is so much longer than the thermal timescale.

The first simulation was run with \( \lambda = 0.5 \) so that the disc should, in general, heat up. The temperature and density are plotted in figures 5.10 and 5.11 respectively. The inner regions of the disc heat up, while the outer regions remain at the same temperature.

The second simulation, with \( \lambda = 1 \) is plotted in figures 5.12 and 5.13. In this case, the inner regions of the disc actually cool, while the outer regions...
still remain the same.

Finally, the third simulation was run with $\lambda = 1.5$, and is plotted in 5.14 and 5.15. In this case, the disc again cools in the inner regions – much more rapidly than with $\lambda = 1$. As the disc particles hit the lowest temperatures, they begin to start heating again. In this case, a temperature equilibrium is not reached.

Using the optically thin cooling law makes the disc much more sensitive to changes in the cooling efficiency than in the optically thick case. In general, these simulations are less numerically stable than their counterparts. This is not particularly surprising, since the density structure, which can’t change on timescales as short as the thermal timescale, was calculated assuming that cooling is optically thick. To illustrate the differences in stability, the timesteps used in simulation Thick2 and Thin3 are plotted together in figure 5.3. Note that these are the timesteps used to alternate between the viscosity and gravity and cooling calculations, not the smaller timesteps used to calculate those steps. The optically thick simulation quickly hits the predefined maximum timestep, while the optically thin simulation has a relatively small timestep, which oscillates rapidly. Over the course of the simulation, the timestep size slowly decreases as the simulation becomes less and less stable.

![Figure 5.3: Timesteps used in the Thick2 and Thin3 simulations. Thick2 is plotted as a solid line, Thin3 is the dashed line.](image)
5.5 Summary

The first sections of this chapter were concerned with the derivation of analytic equations for an azimuthally symmetric thin disc. These equations were derived under a number of assumptions – one of which was that the disc be optically thick and cool as a blackbody. Before performing simulations where this assumption was dropped, it was necessary to first examine the method of modelling the inner boundary. A new method, using ‘buffer’ particles whose density remained constant was proposed. This method resolves the inner boundary well, allowing the modelling of non-isothermal discs.

Finally, six sets of simulations were run, exploring two different cooling laws each with three different levels of cooling efficiency. Those simulations which were optically thick were found to be very insensitive to the cooling efficiency and very quick to regain equilibrium. On the other hand, the simulations with the optically thin cooling law were much more sensitive to the efficiency. They did not regain equilibrium within the time frame of the simulations, even for simulations run up to ten time units, which indicates that they would never reach an equilibrium state. The Strang Splitting code with the error control mechanism described in chapter 3, however, proves to be very robust as it is able to follow the changes in temperature for a long time.

In the next chapter we will turn to accretion discs in cataclysmic variables. The more complicated cooling equations will be dropped, and we will return to modelling isothermal discs.
Figure 5.4: Simulation Thick1: Energy profiles. \( \lambda = 0.1 \). For each time frame, the top plot shows the x-y plane coloured according to energy, the bottom plot shows energy versus radius.
Figure 5.5: Simulation Thick1: Density profiles. $\lambda = 0.1$. For each time frame, the top plot shows the x-y plane coloured according to density, the bottom plot shows density versus radius.
5.5. SUMMARY

Figure 5.6: Simulation Thick2: Energy profiles. $\lambda = 1$. For each time frame, the top plot shows the x-y plane coloured according to energy, the bottom plot shows energy versus radius.
Figure 5.7: Simulation Thick2: Density profiles. $\lambda = 1$. For each time frame, the top plot shows the x-y plane coloured according to density, the bottom plot shows density versus radius.
Figure 5.8: Simulation Thick3: Energy profiles. $\lambda = 10$. For each time frame, the top plot shows the x-y plane coloured according to energy, the bottom plot shows energy versus radius.
Figure 5.9: Simulation Thick3: Density profiles. $\lambda = 10$. For each time frame, the top plot shows the x-y plane coloured according to density, the bottom plot shows density versus radius.
Figure 5.10: Simulation Thin1: Energy profiles. $\lambda = 0.5$. For each time frame, the top plot shows the x-y plane coloured according to energy, the bottom plot shows energy versus radius.
Figure 5.11: Simulation Thin1: Density profiles. $\lambda = 0.5$. For each time frame, the top plot shows the x-y plane coloured according to density, the bottom plot shows density versus radius.
Figure 5.12: Simulation Thin2: Energy profiles. \( \lambda = 1 \). For each time frame, the top plot shows the x-y plane coloured according to energy, the bottom plot shows energy versus radius.
Figure 5.13: Simulation Thin2: Density profiles. \( \lambda = 1 \). For each time frame, the top plot shows the x-y plane coloured according to density, the bottom plot shows density versus radius.
Figure 5.14: Simulation Thin3: Energy profiles. $\lambda = 1.5$. For each time frame, the top plot shows the x-y plane coloured according to energy, the bottom plot shows energy versus radius.
Figure 5.15: Simulation Thin3: Density profiles. $\lambda = 1.5$. For each time frame, the top plot shows the x-y plane coloured according to density, the bottom plot shows density versus radius.
Chapter 6

Cataclysmic Variable Discs

There are approximately four hundred thousand million stars in our galaxy and only about half of these exist alone—the rest are in binary or multiple systems. With that in mind, it is time now to move from examining the simple model of a disc around an isolated object to the more complex situation where a disc accretes onto a star in a binary system. We will first examine the physical structure of these systems, and the observational evidence for that structure. We will then investigate the effects of various binary parameters on the structure of the accretion disc.

6.1 The Basic Model

In the first part of this thesis, in section 2.5, an approximation for the gravitational potential of a test particle in a binary system was introduced. The effective equipotential curves, corotating with the binary, give some insight into the shape and behaviour of a binary system. The lines of constant potential on the orbital plane are plotted in figure 6.1. As stated in section 2.5 it is assumed that the two stars may be considered to be centrally condensed and orbiting in circular orbits with speed $\Omega$. In this figure, the two stars are denoted $m_1$ and $m_2$ (which also denote their masses) and their separation
Figure 6.1: Roche Equipotentials for $m_1 = 2m_2$.

has been scaled to one unit, so that $m_1$ is at the origin and $m_2$ is at (0,1) in the $x-y$ plane. The distance between the two stars is usually called $a$. This figure shows the Roche equipotentials for a system with $m_1 = 2m_2$. The centre of mass is marked with a cross.

At points close to each star, the surfaces of constant potential look like spheres, however as the distance from the stars increases, these surfaces become more and more distorted, forming ellipsoidal shapes with the major axis along the line joining the two stars. The point at which the two ellipsoids just touch is known as the $L_1$ point, or the “inner Lagrangian point”, while the shapes made by the ellipsoids are termed “Roche Lobes”. Throughout this thesis the star $m_1$ will be referred to as the primary and the star $m_2$ as the secondary.

Figure 6.2 shows the surface of the Roche potential of the same two stars. The larger well at the left is around the heavier star. It is clear in this diagram that the inner Lagrangian point, $L_1$, between the two stars is a saddle point. This point is a pass between the two stars. The outer edges of the diagram
curve down because of the centrifugal force.

Binaries may be classified according to how large each star is in relation to its Roche lobe. If the photospheres of both stars lie well within the Roche lobes then their influence on each others' evolution is minimal. These systems are detached binaries. Semi-detached binaries are those in which one of the stars fills its Roche lobe while the other does not, and in contact binary systems both stars fill their Roche lobes.
6.2 Cataclysmic Variables

Cataclysmic variables are semi-detached binaries in which the primary — the star which does not fill its Roche lobe — is a white dwarf. The secondary is typically a low mass \((0.1M_\odot - 1M_\odot)\) star on or near the main sequence. Since the secondary is filling its Roche lobe, there is mass transferred between the two stars. The path this stream of matter takes depends on the ratio of the two stars’ masses and the presence of magnetic fields. In the absence of a strong magnetic field, the stream will generally circle the white dwarf, eventually spreading to form a disc.

Cataclysmic variables are classified into magnetic and non-magnetic systems and then into many subtypes. It is the non-magnetic stars which are the focus of this chapter. They may be further split into stars which have been observed to erupt only once (Classical Novae - CN), stars which have been observed to erupt many times (Recurrent Novae - RN and Dwarf Novae - DN), and stars which have never been observed to erupt (Nova-Like - NL and Novae Remnants - NR) [65]. Of course, stars may move amongst the categories when more is learned about them, or if their behaviour changes. The classification of cataclysmic variables is shown in figure 6.3.

Those stars which erupt many times are further split into two categories, depending on the type of the eruption. Recurrent novae have eruptions in which a substantial shell is ejected at high velocities. These are generally thought to be the result of a thermonuclear runaway on the surface of the white dwarf and, as with classical novae, tend to increase the system’s visual magnitude by a factor of about 8 to 15 magnitudes over a few hours. Dwarf Novae have outbursts which are thought to originate in the disc itself. These eruptions are far less spectacular than RN or CN eruptions, usually only increasing the system’s brightness by about 2 to 5 magnitudes, but are more obviously cyclic, occurring frequently with periods of weeks to months.

The dwarf novae can be further classified on the basis of additional observations:

**U Gem** stars are the archetypal dwarf novae. They show outbursts regularly every few weeks to months, and don’t show any other features.
Figure 6.3: The family tree of cataclysmic variables.

**Cataclysmic Variables**

- **Non-Magnetic**
  - **Number of observed eruptions**
    - **none**
      - **Nova-like Variables**
        - **UX UMa**
        - **RW Tri**
      - **Nova Remnants**
    - **one**
      - **Classical Novae**
  - **many**

- **Magnetic**
  - **Synchronous**
    - **Shell ejected**
      - **yes**
        - **Recurrent Novae**
          - **Z Cam**
        - **no**
          - **Dwarf Novae**
            - **U Gem**
            - **SU UMa**
  - **Asynchronous**
    - **Hard X-Rays**
      - **yes**
        - **Intermediate Polars**
      - **no**
        - **DQ Her**
Some authors use the term U Gem as a synonym of dwarf novae, using the classification SS Cyg to differentiate those stars which are neither Z Cam or SU UMa types [1].

**Z Cam** stars are dwarf novae which occasionally show protracted standstills, in which the brightness remains slightly below outburst maximum for a period of tens of days to years.

**SU Uma** stars show superoutbursts in which the brightness is approximately 0.7 to 1 magnitude brighter than a normal outburst. Superoutbursts last about 5 times as long as normal outbursts, and have an additional modulation known as a superhump which typically has a period a few percent longer than the orbital period.

The nova-like stars are also further classified on the basis of photometry and spectroscopy:

**UX UMa** stars show persistent broad Balmer absorption line spectra, and have also been called 'thick disc' CVs. It is thought that they may be Z Cam stars in extended standstill [40].

**RW Tri** stars show pure emission spectra and, to some extent, may differ from the UX UMa stars only by orbital inclination [76].

**SW Sex** stars may belong to either the UX UMa or RW Tri classes, and show significant phase shifts and transient absorption lines at some phases [73].

**VY Scl** stars have also been termed the 'anti-dwarf novae', and like the SW Sex stars may belong to either the UX UMa or RW Tri classes. They generally vary little about a mean magnitude, but occasionally have a rapid reduction in brightness by more than 1 magnitude, where they remain for weeks to years appearing spectroscopically like the DN at quiescence.

Brian Warner's book *Cataclysmic Variable Stars* [76] is generally thought to be the definitive reference on the subject of cataclysmic variables, their subtypes and observations.
6.3 SPH Simulations of Cataclysmic Variables

SPH is well suited to modelling cataclysmic variables. It is able to adapt to changes in density, to handle open boundaries and non-axisymmetric simulations with ease. These advantages have been utilised by many groups who study accretion discs in cataclysmic variables using SPH. As early as 1989, Molteni and Giannone [43] were evaluating SPH simulations of accretion discs in cataclysmic variables, and found the method to be “quite promising”.

Simulations of the impact between the accretion stream and the disc in three dimensions have shown that substantial amounts of the stream can overflow the disc. Kunze, Speith and Hessman [26] computed high resolution SPH simulations of five different cataclysmic variables, with different mass transfer rates and thermal states. They determined that the accretion stream was not stopped at the edge of the disc – rather, the stream was deflected vertically and travelled across the upper and lower surfaces of the disc to the circularisation radius. In earlier work, Meglicki, Wickramasinghe and Bicknell [39] computed a three dimensional model of a truncated accretion disc such as might be found in an intermediate polar. They also found that the stream overflowed the disc.

The presence of non-axisymmetric spiral shock waves has been noted in many SPH accretion disc simulations, and their cause much debated. Lin, Papaloizou and Savonije [32], [31] and Lubow and Pringle [35] both argued that in three dimensions, spiral shocks would be deflected upwards at the edges of the disc (where they formed), and so would not be seen. Molteni, Belvedere and Lanzafame [42], and Lanzafame, Belvedere and Molteni [27] computed SPH simulations of accretion discs in three dimensions and found that spiral shocks did not form. Furthermore, they found that discs did not form at all when the specific heat ratio in the polytropic equation of state was larger than approximately 1.2. Haraguchi, Boffin and Matsuda [12], however, performed simulations in three dimensions which showed the presence of spiral shocks, as did Yukawa, Boffin and Matsuda [80] – who also showed that discs did form for higher polytropic indices. Lanzafame, Maragvigna and Belvedere [28] performed a set of simulations with varying mass ratio,
and showed that the mass ratio was a dominant factor in the production of spiral shock waves – their results suggested that spiral shock waves should only form for systems with lower mass ratios. Further investigation [29] into the role of the injection velocity determined that an increase in the angular momentum of the incoming gas stream favoured the onset of shocks. Of course, they showed only a correlation between the mass ratio and presence of shocks – which does not imply a causal relationship. Most authors believe that spiral shock features are a result of a tidal deformation of the disc at the outer boundary that is sheared into a spiral pattern by the (near) Keplerian rotation profile of the disc material (see, for example, [5] or [67]).

6.4 Simulations

In order to simulate the disc in a cataclysmic variable, the Roche model must be examined in a little more detail. Variables here are scaled using the scaling relations in section 2.8. Mass is scaled to the total mass of the system and distance to the binary separation.

A binary will be typically characterised by either the mass ratio \( q = m_2/m_1 \), or by the reduced mass of the primary, \( \mu = m_1/(m_1 + m_2) \). The relationship between the two is, of course,

\[
q = \frac{1 - \mu}{\mu}.
\]

I will tend to use the mass ratio in this thesis. The other parameter important in characterising a binary system is the separation, \( a \). With these two parameters most of the details of the Roche geometry may, at least, be estimated with good accuracy. The most important for disc modelling are the location of the \( L_1 \) point and the angle of the stream of matter injected from the secondary.

The distance between the \( L_1 \) point and the centre of the primary star is given by the fitted formula of Plavec and Kratochvil [60]:

\[
\frac{b_1}{a} = 0.500 - 0.227 \log_{10} q
\]
Lubow and Shu [34] calculate the angle that the stream makes with the line joining the centres of the two stars as

$$\cos 2\theta_s = -\frac{4}{3A} + \left(1 - \frac{8}{9A}\right)^{1/2}$$

(6.1)

where

$$A = \frac{\mu}{|b_1 - 1 + \mu|^3} + \frac{1 - \mu}{|b_1 + \mu|^3}$$

or

$$A = \frac{1}{|b_1(q + 1) - q|^3} + \frac{q}{|b_1(q + 1) + 1|^3}$$

Each disc simulation begins with no particles. The mass transfer rate, mass ratio, and binary separation are set. Particles are injected into the simulation through the inner Lagrangian point at a constant rate — regardless of the timestep size — and at an angle to the x-axis defined by equation (6.1). The mass transfer rate is $6 \times 10^{16}$ g/s, and particles are injected with a (scaled) velocity of 0.02, which corresponds to the thermal velocity in the upper atmosphere of the secondary star.

The inner boundary condition is that described in section 5.2. A particle is deemed to have been accreted once it falls over the inner radius, $r_{\text{in}}$. In order to prevent particles at opposite sides of the hole from sensing each other, this is set to be slightly larger than twice the maximum possible smoothing length.

Ideally, simulations should be run at least until the accretion rate matches the injection rate — that is, until the number of particles in the disc is constant. An estimate of how long this should take can be obtained using the results of thin disc theory, even though the theory itself won’t necessarily hold. Equation (5.16) gives a relation for the surface density of the disc:

$$2\pi a \Sigma c_s^2 = \dot{M} \Omega \left(1 - \left(\frac{r_0}{r}\right)^{\frac{3}{2}}\right)$$

(5.16)
which we can integrate over the entire surface to obtain the mass of the disc. Dropping numerical constants,

\[ M_{\text{disc}} \propto \frac{\sqrt{m_1 \dot{m}}}{\alpha}. \]

This means that with a large viscosity the mass of the disc will be small, which in turn means that a smaller number of particles will be needed to simulate the disc. From equation (2.14)

\[ \alpha H = \frac{1}{8} \bar{\alpha} h \]  \hspace{1cm} (2.14)

We have no way of knowing a priori what the ratio of the smoothing length to the scale height of the disc will be – indeed, it will vary with radius (which simply highlights that neither the Shakura–Sunyaev \( \alpha \) nor the SPH \( \bar{\alpha} \) can be assumed to be constant, even though they often are for convenience). As a rough guide, however, we can assume that \( 0.01 \leq h/H < 1 \), so a physical range of \( \bar{\alpha} \) will be \( 8\alpha < \bar{\alpha} < 800\alpha \), which is quite a large range to play with. Nonetheless, we know that \( \alpha < 1 \) and probably \( \alpha \ll 1 \), so with \( \alpha = 0.1 \), we get \( 0.8 < \bar{\alpha} < 80 \). With \( \alpha = 0.01 \), we get \( 0.08 < \bar{\alpha} < 8 \). Of course, since the mass of the disc is inversely proportional to \( \bar{\alpha} \), as \( \bar{\alpha} \) is increased the time taken to evolve the simulation to a steady state is decreased. For the following section, \( \bar{\alpha} = 2 \) has been chosen. It is well within the range of reasonable values, and is large enough that the time taken to run the simulations is not too long. The non-linear viscosity parameter, \( \beta \) was set to be \( \beta = 1 \). This is enough to resolve the shock where the stream meets the disc, and to stop the stream impinging on the disc. Simulations performed in this chapter and the next assume an isothermal equation of state.

### 6.5 Varying the mass ratio

The simulations here examine the most central of binary parameters, the mass ratio. Set A has a mass ratio of \( q = 0.2 \) while set B has \( q = 0.8 \). The first thing to examine is how the disc builds up. The first few orbital periods
VARYING THE MASS RATIO

for set A are plotted in figure 6.4 and for set B in 6.5. Each plot shows the particle positions as dots. The solid line is the Roche Lobe for the primary, and the primary itself is plotted as a small circle at the origin.

In both sets, the stream initially passes very close to the primary on the opposite side to the $L1$ point. The distance, $r_{\text{min}}$, that the stream has from the centre of the primary at closest approach is given by Lubow and Shu [34] as

$$r_{\text{min}} = 0.0488q^{-0.464} \quad 0.05 < q < 1$$

(6.2)

which gives $r_{\text{min}} = 0.103$ for $q = 0.2$ and $r_{\text{min}} = 0.054$ for $q = 0.8$. The stream then loops back and intersects itself relatively close to the binary, where the shock of impact will dissipate much of the kinetic energy, but not the angular momentum. This then forms a thick ring of material which, to begin with, is slightly elongated but quickly settles to a circular ring located at the radius where the angular momentum of a Keplerian orbit is equivalent to the angular momentum of the incoming stream. This circularisation radius is approximated by Hessman and Hop [18], taking into account the effect of the secondary removing angular momentum from the stream, as

$$r_{\text{circ}} = 0.0859q^{-0.426} \quad 0.05 < q < 1.$$ 

(6.3)

For $q = 0.2$, this gives $r_{\text{circ}} = 0.1705$ and for $q = 0.8$, $r_{\text{circ}} = 0.094$. A quick check with figures 6.4 and 6.5 shows that both the circularisation radius and the distance of closest approach obtained in simulations A and B are in good agreement with the values obtained theoretically.

Due to the viscous dissipation, the ring begins to spread out in much the same way as the ring studied in section 4.2. At this stage, there is little difference between the two sets of simulations. Set A, with a lower mass ratio, has a larger Roche Lobe, and a larger ring – as is to be expected from equation 6.3.

The longer term spatial evolution of the simulations in set A is shown in figure 6.6, while that of set B is in figure 6.7. By about the first 120 orbital periods, the spatial extent of the disc has reached its outer limit in both cases, even though the number of particles is still quite low. From here,
Figure 6.4: Simulation A: building the disc with $q=0.2$
6.5. VARYING THE MASS RATIO

Figure 6.5: Simulation B: building the disc with $q=0.8$
the additional particles begin to thicken the high density ridge around the
rim of the disc, and extend to fill in the centre regions. This outer ridge
corresponds to the largest simple periodic orbit which does not intersect any
other orbit. These are identified in Paczyński [58] and are typically thought
of as the outer edge of the disc, although the effect of viscosity and pressure
may push the true disc edge somewhat past this radius.

Due to time constraints, these simulations were not run until the particle
number reached equilibrium. This is evident in the low resolution in the
inner regions of the disc. The outer regions, however, are well resolved, and
the general shape and behaviour of the disc should not be greatly effected
by the lack of resolution in the inner regions.

At the time of analysis, the number of particles in simulation A was over
41,000, while simulation B had over 21,000 particles – the particle number
was still increasing. In each simulation, particles had between 60 and 100
neighbours. Those particles which lay in the inner, under resolved region
were at the lower end of the range. The majority of particles had between
80 and 90 neighbours.

**Disc Precession**

At approximately 200 orbital periods, the disc in simulation A begins to
precess. Disc precession was proposed to be the source of the superhumps
found in SU UMa stars during superoutburst by Osaki [55] in 1985. In 1988,
Whitehurst ([78] and [79]) produced numerical simulations of the system Z
Cha ($q = 0.15$), showing a disc slowly precessing after a single mass transfer
burst. Hirose and Osaki [19] simulated accretion discs for a range of mass
ratios, finding that precession occured in discs in systems with $0.05 < q < 0.25$.

A particle at the outer edge of the accretion disc will be able to “see” the
secondary each time it moves past on its own orbit. At this time, the particle
will recieve a “kick” due to the gravitational force of the secondary, which
will slightly increase the eccentricity of the particle’s orbit. If it happens
that the secondary always moves past the particle at the same point in the
Figure 6.6: The longer term evolution of the disc with $q=0.2$
Figure 6.7: The longer term evolution of the disc with $q=0.8$
6.5. VARYING THE MASS RATIO

particle’s orbit, then the eccentricity of the particle’s orbit will grow. If \( \Omega \) is the particle’s mean angular frequency around the primary in the stationary frame and \( \omega \) is the apsidal precession frequency of its orbit, then the epicyclic frequency is \( \Omega - \omega \). From the point of view of the particle, the secondary moves with frequency \( \Omega - \Omega_{\text{orb}} \), and so for the eccentricity of the particle’s orbit to grow, we require this frequency to be an integer multiple of the epicyclic frequency:

\[
k(\Omega - \omega) = j(\Omega - \Omega_{\text{orb}})
\]

where \( j \) and \( k \) are positive integers.

An estimate of the size of a resonant orbit for a given \( j \) and \( k \) can be obtained by assuming that the mean radius of the particle’s orbit is close to the Keplerian value, which gives

\[
R_{jk} = \left(\frac{j - k}{j}\right)^{2/3} (1 + q)^{-1/3}
\]

where \( R_{jk} \) is in terms of the binary separation. Comparing the values of \( R_{jk} \) for various choices of \( j \) and \( k \) with an estimate of the disc radius tells us whether or not a disc in a given system will extend far enough out to begin to precess. One such estimate is the family of largest simple periodic orbits calculated by Paczyński [58]. These orbits are elongated in the y-direction, and not symmetric in the x-direction and so there are three parameters which describe their radius. \( r_{\text{max}} \) is the maximum radius of each orbit, \( r_1 \) is the distance from the primary to the intersection of the orbit and the x axis on the side furthest from the secondary, \( r_2 \) the distance from the primary to the intersection of the orbit and the x axis on the side closest to the secondary. Another estimate of the size of the disc is simply to restrict it to some fraction of the mean Roche Lobe size. This is most commonly stated as

\[
R_{\text{trunc}} = 0.9 R_1
\]

where \( R_{\text{trunc}} \) is the truncation radius of the disc.

These four disc radii are plotted in figure 6.8 along with several resonant radii. The \( j = 3, k = 2 \) resonance is the only one which occurs at a radius
small enough for a cataclysmic variable disc to reach, and then only when \( q \lesssim 0.33 \) according to \( R_{\text{trunc}} \) or \( q \lesssim 0.4 \) using Paczyński’s \( r_{\text{max}} \). Murray [50] calculated the rate of dynamical precession as

\[
\omega_{\text{dyn}} = a(r) \frac{q}{\sqrt{1 + q}} \Omega_{\text{orb}}
\]

with

\[
a(r) = \frac{1}{4r^{1/2}} \frac{d}{dr} \left( r^2 \frac{d b^{(j)}_{1/2}}{dr} \right).
\]

\( b^{(j)}_q \) is the standard notation for a Laplace coefficient from celestial mechanics. Murray has plotted the function \( a(r) \), and for values of \( r \) at the \( j = 3, k = 2 \) resonance for a disc in a \( q = 0.2 \) system, \( a(r) \approx 0.4 \), which gives a precession period of \( P_{pr} \approx 1.0788 P_{\text{orb}} \).

For the case in simulation A, \( q = 0.2 \), the extent of the disc is certainly large enough that the outer regions of the disc lie within the resonance radius. This introduces a tidal instability in the outer regions of the disc, even before the disc has reached a state where the mass transfer rate is matched by the accretion rate. The disc is shown coloured according to density in figure 6.9. The density is interpolated onto a grid using the same procedure used for the figures shown at the end of chapter 5. It is displayed on a logarithmic scale in order to highlight the structure. These figures cover a span of approximately 1.2 orbital periods – slightly longer than the period of the precession.

Figure 6.10 shows the viscous dissipation at the same times as the density plots. Most of the dissipation comes from the centre regions of the disc where the particles are moving quite rapidly, but are not affected by the precession of the outer regions. In order to highlight the detail of the edge structure in the plots of both dissipation and density, I have blanked out the inner regions.

Examining the density and dissipation plots reveals that the high density ridge is not constant over the period of the precession. The outer parts of the disc move on eccentric orbits and do not precess as a rigid body. As the particles move relative to each other the ridge seems to alternate between being sharply in focus and being quite diffuse.
In contrast, the coloured density plots for simulation B, shown in figure 6.11 are much more uniform. In each time frame, the disc has a flatter density profile – the high density ridge is present, but not as prominent as it was in simulation A. Across the time frames, there is very little variation in the density structure. The high density ridge appears much wider in these images than the corresponding ridge in simulation A, however this is simply due to the scale.

The viscous dissipation, shown in figure 6.12, highlights the ridge a little more than the density plots, however the point where the stream impacts the disc is definitely the most prominent feature. Again, there is very little variation across the time frames.

The precession period can be obtained exactly by examining the viscous dissipation in the disc. At each output time, each particle’s viscous dissipation is also output. The dissipation can then be summed over the whole disc or specific regions of the disc. A Lomb normalised periodogram was performed using the LNP\_TEST routine of IDL, which is based on fasper from Numerical Recipes [62]. This routine returns the spectral power as a function of linear frequency. In figure 6.13 this spectral power function is plotted for four different parts of the disc. First, the dissipation over time for the whole disc was analysed. This produced a large number of broad, insignificant peaks. Excluding the very inner regions of the disc produces small peaks at \( f = 0.144 \) and \( f = 0.288 \), as well as leaving a set of weak peaks spread across the spectrum. Excluding more, and then even more, of the inner disc, shown in the third and fourth plots of figure 6.13, strengthens the peak at \( f = 0.144 \) and reduces everything except for the harmonics of this peak to the level of noise. This peak corresponds to a period of \( 1.10P_{\text{orb}} \), which is in excellent agreement with the theoretical precession rate calculated by Murray [50].

I have also used LNP\_TEST to search for periodicities in the viscous dissipation of simulation B. The results are shown in figure 6.14. In this case, even the largest peak at \( f = 0.035 \) for the outer parts of the disc is below noise levels. In this disc there is no precession and no periodic variation in the viscous dissipation.
Figure 6.8: Relevant disc radii with discs in cataclysmic variables.
Figure 6.9: Simulation A: density plot for $q=0.2$. See text for details.
Figure 6.10: Simulation A: dissipation plot for $q=0.2$. See text for details.
Figure 6.11: Simulation B: density plot for q=0.8. See text for details.
Figure 6.12: Simulation B: dissipation plot for $q=0.8$. See text for details.
Figure 6.13: Lombic periodogram of the dissipation from the precessing disc. x axis is frequency, y axis is spectral power.
Figure 6.14: Lombic periodogram of the dissipation from the non-precessing disc. x axis is frequency, y axis is spectral power.
6.6 The Viscosity Parameters

It is worthwhile to briefly examine the effects of the chosen viscosity parameters on the final form of the disc. To do this, I have run an additional simulation with $\alpha = 10$ and $\beta = 0$. This large value of $\alpha$ ensures that the disc evolves quickly, although the final particle number will be low. At the end of the simulation run, the total particle number was approximately 12,000 – it was essentially steady. The number of neighbours of each particles was between 80 and 100. This simulation is performed with $q = 0.2$. It will be referred to as simulation C, and is comparable to simulation A, which was described in sections 6.4 to 6.5. All other parameters are identical to simulation A.

The initial development of the disc is shown in figure 6.15. This figure is analogous to figure 6.4. Although there are some differences, the main features, the location of the circularisation radius and the point where the stream intersects itself, are the same. The ring is broader in simulation C, which is most likely due to the larger shear viscosity.

Figure 6.16 shows the development of the disc of simulation C at later times – it is analogous to figure 6.6. There are several important differences between these two plots. Initially in figure 6.16 the disc is much broader. It spreads out more quickly, reaching the last non-intersecting orbit by about 20 orbital periods – compared to around 120 orbital periods in figure 6.6. In both cases the high density ridge around the outside of the disc is obvious, however figure 6.16 also shows a thickened ring at around the circularisation radius. This is most likely because the shock where the stream hits the disc is not resolved, and stream particles are allowed to impinge on the disc. On the other hand, the central regions of simulation A are under resolved, even after a much longer evolution.

The density and dissipation of the disc of simulation C over one precession period are shown in figures 6.17 and 6.18. These plots are analogous to 6.9 and 6.10 respectively. The shape of the disc as it precesses is much the same. The obvious difference here is visible in the dissipation plots. In figure 6.18 the stream is seen impinging on the disc as a bright, hot, dagger-like shape.
In figure 6.10, however, the hot stream impact region is confined to the edge of the disc.

Finally, figure 6.19 shows the lombic periodogram of the disc with $\alpha = 10$ and $\beta = 0$. As with figure 6.13, first the whole disc is analysed, then the inner regions are excluded. The peaks in the latter case are far broader (note also that the x-axis scale is different), however they are in the same position. Essentially, both of these simulations give the same precession period at the same strength — regardless of the difference in viscosity parameters.

6.7 Summary

This chapter began with a brief review of cataclysmic variables and the Roche Lobe model for the gravitational potential in a binary system. Two examples of cataclysmic variables were simulated — one with a mass ratio of $q = 0.2$ and one with $q = 0.8$. It didn’t take long for the $q = 0.2$ disc to begin precessing. The period of disc precession was determined to be 1.1 times the orbital period, in excellent agreement with theoretical predictions. A further simulation was calculated in which the viscous parameters were changed. Although some features of the disc changed, the general shape and precession did not.

In the next chapter, a specific example of a cataclysmic variable is discussed — one in which the mass ratio is thought to lie between 0.62 and 1.04. According to theoretic results, and the simulations presented in this chapter, a disc simulation with a mass ratio in that range will not show precession. Chapter 7 will explore one possible scenario which may cause that disc to precess.
6.7. SUMMARY

Figure 6.15: Building the disc with $q=0.2$, $\bar{\alpha} = 10$ and $\beta = 0$. 
Figure 6.16: The further evolution of the disc with $q=0.2$, $\alpha = 10$ and $\beta = 0$. 
Figure 6.17: Density plot for $q=0.2$, $\alpha = 10$ and $\beta = 0$. 
Figure 6.18: Dissipation plot for $q=0.2$, $\alpha = 10$ and $\beta = 0$. 
Figure 6.19: Lombic periodogram of the dissipation from the disc with $q=0.2$, $\alpha = 10$ and $\beta = 0$. x axis is frequency, y axis is spectral power.
Chapter 7

TV Columbae – A Hunt for Precession

Up to now, we have assumed that the inner boundary is very close to the central object. In fact, we have tried to put it as close as possible, in order to minimise the effects of attempts to model it. In this chapter, a system is examined in which the inner boundary is probably not close to the primary. The system in question is TV Columbae – an intermediate polar. The intermediate polars are a subclass of cataclysmic variables where the primary’s magnetic field is strong enough to disrupt the inner regions of the accretion disc, but not strong enough to lock the secondary into a synchronous orbit. Intermediate polars are multi-periodic and so are a rich source of information about the physical processes involved in accretion and in cataclysmic variables in general [76].

7.1 TV Col

TV Columbae was first noted as an x-ray source in the 1978 Ariel V survey [8] and named 2A 0526–328. Charles et al [7] examined its emission line spectrum and found it to be similar to AM Her, although searches for polar-
ization found nothing. They therefore proposed that 2A 0526–328 may be a Z Cam type system in a very extended standstill or a recurrent nova which had not erupted recently. Photometric and spectroscopic studies confirmed the classification as a CV type star, and found three periods: a spectroscopic period at 5.5 hours; a photometric period at 5.2 hours; and a beat between the two at 4.024 days ([21] and [49]). Hutchings et al [21] suggested that the spectroscopic period is associated with the orbital motion of the binary and the photometric period with the rotation of the primary. They proposed that the mechanism giving rise to the multiple periodicities observed was the non-synchronized rotation of a weakly magnetic white dwarf, which rotates once in the binary frame every 4.024 days. However, x-ray observations in 1985 and 1987 ([69],[68]) found a pulsation period of 1911 seconds (0.53 hours). This was determined to be the true rotation period of the white dwarf, while observations of eclipses [16] confirmed the binary orbital period to be 5.5 hours. More recent observations by Retter et al [64] have found yet another period at 6.3 hours. I have summarised the periods of TV Col and their nature in table 7.1.

Table 7.1: The periods of TV Col

<table>
<thead>
<tr>
<th>Period [h]</th>
<th>Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5307</td>
<td>Spin</td>
</tr>
<tr>
<td>5.2</td>
<td>Negative Superhump</td>
</tr>
<tr>
<td>5.4864</td>
<td>Orbital</td>
</tr>
<tr>
<td>6.3336</td>
<td>Positive Superhump</td>
</tr>
<tr>
<td>96.624</td>
<td>Nodal Precession</td>
</tr>
</tbody>
</table>

Probably the most interesting aspects of TV Col are its superhump periods at 5.2 and 6.3 hours. Superhumps are thought to be caused by an eccentric disc which is precessing at a rate either slightly above or below the orbital speed [55]. The observed superhump period $P_{SH}$ corresponds to the period of disc precession $P_D$ in the frame of the binary.

Superhumps have been observed in a large number of systems, but in general, aside from a few exceptions, only in systems with $q \lesssim 0.33$. The difference between this limit, and that found by Hirose and Osaki [19] ($q < 0.25$) is probably related to the numerical method used by Hirose and Osaki – they used the
7.2. SIMULATION DETAILS

Theoretically, it is only in systems with a mass ratio below this limit that the disc extends close enough to the resonance radius for coherent precession to occur, which is what makes the superhumps of TV Col so interesting. Hellier [15] concluded from a spectroscopic analysis that the mass ratio was $q = 0.62 - 0.93$, however Retter et al [64] used equation (15) of Osaki [56] and found a mass ratio of $q = 0.92 \pm 0.12$, which is consistent with the result from Hellier, and well above the theoretic limit for superhumps.

In chapter 6 the different radii involved in the excitation, or not, of disc precession were discussed. There were two measures of possible disc size presented – the tidal truncation radius, $T_{\text{trunc}} = 0.9R_1$, and the last simple periodic orbit, $r_{\text{max}}$. Looking again at figure 6.8, this time focussing on a mass ratio of around 0.6 - 0.9 we see why TV Col theoretically shouldn’t show superhumps: the disc radius, using either measure, is not as large as the smallest resonance radius. It is possible, however, that the magnetic field of the primary could play a role in the size of the disc. One effect the magnetic field could have is at the inner boundary.

When the disc material threads onto the magnetic field lines, it is uncertain what happens to its angular momentum. It is possible that the gas only takes some of its angular momentum with it, leaving the remainder behind in the disc. This excess angular momentum deposited in the centre of the disc could certainly effect the rest of the disc – in particular it should effect the radial extent of the disc. If this is the case, it may explain the presence of superhumps in TV Col. If the disc is forced to be larger than theory predicts, it could extend out as far as the resonance radius and then begin to precess.

7.2 Simulation Details

In order to test this proposition, I have modelled the accretion disc in TV Col with a different condition at the inner boundary. In these simulations, each accreting particle leaves a fraction of its angular momentum in the disc. This is accomplished by first determining the neighbours of an accreting particle

viscous particle method of Lin and Pringle [33] which doesn’t include pressure forces.
and how far away they are. The angular momentum to be deposited in the
disc is divided up amongst the neighbours such that each neighbour gets an
amount proportional to the value of the smoothing kernel at their distance to
the accretor. The accreted particle is accreted directly – the buffer particle
method of the previous chapters is not used.

I have simulated a set of accretion discs with a mass ratio of 0.9, which
is consistent with the analyses of both Hellier and Retter et al. Each simula-
tion was run with the boundary condition described above, with a different
fraction of angular momentum left behind. That fraction is referred to as
“\( \text{jfrac} \)”. The simulation names and their values of \( \text{jfrac} \) are shown in table
7.2. These values were chosen to cover the range of possible values between
0 and 1.

The location of the inner boundary is chosen somewhat arbitrarily. In
theory, we should be able to use the strength of the magnetic field of the
primary to obtain a reasonable estimate, either by using the Alfvèn radius,
or by including the magnetic tension and pressure, and the effects of the spin
of the primary [38]. Unfortunately, a measure of the magnetic field strength
of the primary of TV Col does not exist. We do know, however, that it
must be strong enough to disrupt accretion in the inner disc, but not strong
enough to synchronise the primary and secondary orbits, which gives a range
of reasonable values. The inner radius chosen here was 0.04 (scaled so that
the binary separation is equal to one) – further investigations could examine
other choices of inner radii.

The simulations in this chapter are run with a mass accretion rate of
$6 \times 10^{16}$ g/s, and particles are injected at the velocity which corresponds to the thermal speed in the upper atmosphere of the secondary. As in simulation B of the previous chapter, each simulation had over 20,000 particles, and each particle had between 60 and 100 neighbours. The majority of particles had between 80 and 90 neighbours.

### 7.3 Results

Colour maps of the density and viscous dissipation from simulation A are shown in figures 7.7 and 7.8 respectively. These are produced using the same method as was used to produce the colour maps in chapter 6. In each plot, x and y axes represent the x and y axes of the simulation. A scale is given above each plot. The solid white line is the Roche lobe. They do not show the spatial variations which are typical of disc precession, as seen in the $q = 0.2$ simulation from chapter 6. Instead, they are much more like the normal disc shown in the $q = 0.8$ simulations. Analogous plots for simulation C are shown in figures 7.9 and 7.10, and for simulation F in 7.11 and 7.12. Again, they don’t show the features which would indicate precession. The remainder of the simulations look the same.

Unfortunately, the hunt for precession was unsuccessful. Not one of these disc simulations showed precession. In order to determine whether or not the disc precessed, I have run the same LNP\_TEST routine as described in chapter 6. This routine should not only determine the likelihood of any periodicities in the viscous dissipation from the disc, but also specify exactly what those periodicities are. The results are plotted in figures 7.1 to 7.6. In each plot, the y axis is the spectral power, and the x axis is frequency. As in the previous chapter, first the power spectrum resulting from examining the whole disc is plotted, then from the disc at radii greater than 0.1, greater than 0.2 and greater than 0.3. In each case, there are no significant peaks – a power level of 10 or more would indicate a significant result.

Additionally, there appears to be little difference between each of the simulations. That is, the percentage of angular momentum deposited at the inner boundary has little effect on the disc.
7.4 Summary

TV Columbae is a fascinating object. It shows two superhumps simultaneously — one at a period shorter than the orbital period and one at a period longer than the orbital period — even though the mass ratio is much larger than that of a cataclysmic variable which would normally have superhumps. One possible cause of these superhumps is that the magnetic field of the white dwarf, while not strong enough to synchronize the spins of the primary and secondary, is strong enough to effect the accretion disc. At the inner boundary of the disc, where the effect of the magnetic field is strongest, plasma being threaded onto field lines may leave behind some or all of its angular momentum. In order to test whether or not this would effect the size of the disc enough to cause superhumps, seven sets of simulations were performed. These simulations had a mass ratio appropriate to TV Col, and a boundary condition whereby the particles being accreted left some fraction of their angular momentum behind in the disc. Each simulation showed similar results. There was no disc precession seen.

The hunt for precession does not end here, however. There are any number of other processes that could be responsible for precession in the accretion disc of TV Columbae. It is possible that the secondary star may have a significant magnetic field, and this field could effect the disc. The most likely effect of the secondary’s magnetic field would be to cause the disc to warp, so that it no longer lies in the orbital plane. This warp may then move around the disc. Murray et al [51] suggest that such a warped disc is the most likely cause of negative superhumps in cataclysmic variables.

Other possibilities that could be investigated include different temperature structures and cooling laws, different mass ratios, different inner radii, and other three dimensional effects. The magnetic field of the primary star may also have further influence on the disc. The magnetic field is anchored to the star, and so at the radius at which the magnetic field strength balances the pressure in the disc, the field lines are rotating at a different rate to the Kepler rotation rate. This could cause disc material to be spun up before it is accreted — rather than receieving angular momentum from the disc, the
primary could actually be a source of angular momentum for the disc.
Figure 7.1: Power spectrum, Simulation A. x axis is frequency, y axis is spectral power
Figure 7.2: Power spectrum, Simulation B. x axis is frequency, y axis is spectral power.
Figure 7.3: Power spectrum, Simulation C. x axis is frequency, y axis is spectral power
Figure 7.4: Power spectrum, Simulation D. x axis is frequency, y axis is spectral power
Figure 7.5: Power spectrum, Simulation E. x axis is frequency, y axis is spectral power
Figure 7.6: Power spectrum, Simulation F. x axis is frequency, y axis is spectral power
Figure 7.7: Simulation A: Density. xy plane coloured according to density.
Figure 7.8: Simulation A: Dissipation. xy plane coloured according to dissipation.
Figure 7.9: Simulation C: Density. xy plane coloured according to density.
Figure 7.10: Simulation C: Dissipation. xy plane coloured according to dissipation.
Figure 7.11: Simulation F: Density. xy plane coloured according to density.
Figure 7.12: Simulation F: Dissipation. xy plane coloured according to dissipation.
Chapter 8

Conclusions and Further Work

The first part of this thesis examined the theory, algorithms and methods behind a Smoothed Particle Hydrodynamics code.

The first chapter was primarily concerned with the mathematical basis of SPH, and with translating the fluid equations of motion into SPH form. This gave a way of finding the values of a function and its derivatives at a set of disordered points. Two methods commonly used to make this process more efficient were also discussed.

In chapter 3 two techniques for stepping the SPH equations forward in time were examined. The first of these was the commonly used Leapfrog predictor–corrector algorithm, which also belongs to the family of Runge-Kutta algorithms. Although this method is adequate, it is far from the most efficient way of integrating the fluid equations. The second method, the Strang splitting algorithm, takes into account the time schemes involved in each equation. This algorithm involves splitting each equation into two (or more) parts and integrating each separately, alternating between the two parts at a time step based on accuracy and stability concerns. It was proven that this alternation retains at least the accuracy of the individual integration methods. A method for choosing an alternation time step which adapts to current conditions was also presented.

The final chapter in part I presented several of the tests performed on the code. These showed that the code was able to capture shocks to good
accuracy, and model the viscous shearing of a ring of particles. The final test performed demonstrated the longer term ability of the code to accurately model a disc in equilibrium.

The second part of this thesis used the code developed in part I and applied it to several different accretion disc scenarios. In chapter 5 the code was used to model azimuthally symmetric thin discs. The theory of thin discs was discussed and analytic equations for steady thin discs were derived. These equations were derived under a number of assumptions – one of which was that the disc be optically thick and cool as a blackbody. Before performing simulations where this assumption was dropped, it was necessary to first examine the method of modelling the inner boundary. A new method, using ‘buffer’ particles whose density remained constant was proposed. Simulations were performed of two kinds of accretion discs – one which was optically thin and one which was optically thick. The optically thin simulations were found to be much more sensitive to the efficiency of cooling, and far less stable than their optically thick counterparts, however, the code was able to follow their evolution for quite some time.

Chapter 6 opened with a brief review of cataclysmic variables and their subclasses. Two examples of cataclysmic variables were simulated – one with a mass ratio of $q = 0.2$ and one with $q = 0.8$. The disc simulated with $q = 0.2$ quickly began to precess. The causes of disc precession in cataclysmic variables were discussed, and the precession period of the simulated accretion disc was found to match very well with the theoretically predicted period.

Finally, chapter 7 examined a specific system. TV Columbae shows two superhumps simultaneously – one positive superhump and one negative superhump – even though the mass ratio is much larger than that of a cataclysmic variable which would normally have superhumps. A possible cause of these superhumps was proposed, and a number of simulations run to test that theory. None of the simulated discs, however, showed precession.

Throughout this thesis, a robust and stable code was developed for fluid modelling. Although here it has only been applied to accretion discs in cataclysmic variables, it could also be applied to just about any other fluid system where self gravity is unimportant. The Strang splitting method and error
control technique could also be applied to numerical computations which don't use SPH.

The code developed here was shown to be robust enough to handle simulations with an unstable cooling law. Only two, very simplistic, cooling laws were tested and so there remains much work to be done in this area. Ultimately, a more realistic cooling law, or even a radiative transfer code, could be incorporated into the SPH code.

Of course, much more work remains to be done on TV Columbæ. The inner boundary condition tested here was a simple one – in reality, what happens at the inner edge of the accretion disc will be much more complex. For instance, this boundary condition only affects the particles as they are accreted, but the effect of the magnetic field should be felt at much larger radii than the accretion radius. This is the subject of an ongoing investigation.
Bibliography


