The University of Reading

Four Dimensional Variational Data Assimilation for Hamiltonian Problems

Laura Watkinson

This thesis is submitted for the degree of
Doctor of Philosophy

Department of Mathematics

September 2006

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Laura Watkinson

Abstract

In this thesis we bring together two areas of mathematics; Hamiltonian dynamics and data assimilation. We construct a four dimensional variational (4d Var) data assimilation scheme for two Hamiltonian systems. This is to reflect the Hamiltonian behaviour observed in the atmosphere. We know, for example, that potential vorticity is conserved in atmospheric models. However, current data assimilation schemes do not explicitly include such physical relationships.

In this thesis, by considering the two and three body problems, we demonstrate how such characteristic behaviour can be included in the data assimilation schemes. In our 4d Var schemes we add a weak constraint that imposes the conservation of the Hamiltonian, the total energy, at the initial time. This is effectively imposing an energy constraint from one data assimilation window to the next.

Our results imply that these weak constraints affect the underlying geometry of the resulting data assimilation solution. We also demonstrate that this constraint reduces the error on this solution and the forecast. By imposing this constraint we are including additional information to the system. Due to the additional term in the cost function gradient, the analysis can only change in such a way as to satisfy this weak constraint.

This thesis therefore demonstrates that the inclusion of similar weak constraints, perhaps using the conservation of potential vorticity, could improve the analysis and forecast for atmospheric models.

Acknowledgements

I would like to begin by thanking my supervisor Prof. Nancy Nichols for her guidance and support, without which I could not have completed this research. I am also grateful to Dr Amos Lawless for his patience and assistance with all aspects of my work. I would also like to thank my Met Office supervisors Dr Ian Roulstone and Prof. Mike Cullen for their advice and direction over the past four years. Thanks are also due to the staff of the Department of Mathematics, particularly to Prof. Mike Baines for his useful suggestions and to Sue Davis for her friendship and encouragement throughout my time at the university.

I acknowledge the financial support of the Engineering and Physical Sciences Research Council (E.P.S.R.C.) and the Met.Office.

To all my friends from the Mathematics department and elswhere, thank you for your patience and support. I am particularly grateful to Claire, Davinia and Martin, I could not have done this without your friendship and moral support. I also wish to thank my parents for their help throughout my years of education; I would not have achieved this much without your love and encouragement. Finally, and with love, I thank my husband Luke; only with your support, faith and love have I been able to complete this PhD.

Contents

List of Figures

Chapter 1

Introduction

In this thesis we consider the development of data assimilation schemes for two Hamiltonian problems. This is to reflect the Hamiltonian behaviour observed in the atmosphere. We intend to demonstrate that the inherent characteristics of such systems can be used explicitly within the data assimilation algorithms. We begin by considering the motivation both for the use of data assimilation schemes in general and for our investigation into Hamiltonian problems.

1.1 Motivation

The atmosphere is a large and chaotic system [33] and therefore its evolution can be difficult to predict. A range of model equations and numerical schemes have been developed to try and solve this problem. In all cases however, there exists the difficulty of choosing initial data. We consider a global model such as the Unified Model at the Met Office, which has a horizontal resolution of $0.8\dot{3}^\circ$ longitude by $0.\dot{5}^\circ$ latitude and 38 vertical levels. This results in a three dimensional grid consisting of approximately $5\text{x}10^6$ points [38]. In addition the model evolves several different variables (of order 10), thus the size of the state vector is of order 10^7. However it is not possible to obtain observations occurring at the same time with sufficient accuracy, density and spatial resolution to initialise such a large model [31]. We now consider techniques that can be used to overcome this problem.

If we consider the atmosphere over all spatial resolutions it is indeed chaotic in nature. However the large scale behaviour is typically slowly varying. This implies that the forces involved are approximately balanced. It is thus possible to reduce the size of the model state by making use of these relationships to constrain the model state vector [31]. These are known as balanced models. Examples of these can be found in [11, 37]. Such models eliminate the fast oscillations that can be spuriously generated by unbalanced models. These small scale, fast oscillations are often referred to as gravity waves. As approximations have been made, these balanced models have limited accuracy. However they often have useful mathematical properties. Semigeostrophic theory, a well known example, has amongst other properties, a Hamiltonian structure. Such characteristics allow for robust numerical methods and in addition give a more complete picture of the mathematical properties of the system [37].

Although these balanced models reduce the size of the state vector, the problem still remains underdetermined. The number of observations is still insufficient to describe the problem and may have limited accuracy. This lack of *completeness* in the initial conditions will result in a poor description of the large scale structures and will thus provide a poor forecast of the low frequency, synoptic waves [31]. Thus it is necessary to find a technique to provide initial conditions for the model in such a way that compensates for this missing information. In addition, if a balanced model is not used, the inaccuracies in the observations may produce spurious gravity waves.

Data assimilation methods attempt to account for this by incorporating observational data into a numerical model in order to provide a best estimate of the initial conditions [5, 13]. However as there are insufficient observations, information from previous forecasts can also be included. This then provides initial conditions for the full state vector thus helping to solve the completeness problem. It may not however completely remove difficulties associated with lack of accuracy and, if the model allows, spurious gravity waves may still be generated.

Techniques have been developed to try and reduce this problem. One such method involves the decomposition of the initial conditions into the fast and slow parts of the motion and then initialising the model using only the slow motion. A second method involves the inclusion of additional constraints in the data assimilation scheme which have the effect of filtering out the spurious fast waves.

In this thesis we consider the technique of *variational* data assimilation. This involves the minimisation of a cost function that measures the distance between a 'first guess' trajectory and the observations and background. In this way we find a complete set of initial conditions to describe the evolution of the system. This assimilation technique can be extended to include further constraints, such as those mentioned above.

As we have discussed, the semigeostrophic model has a *Hamiltonian* structure. It has also been shown that higher order balanced models for the atmosphere are also Hamiltonian [37]. It is observed that the atmosphere can be considered Hamiltonian over the period of one to three or four days as the external forcings acting over this timescale are negligible. Such systems have mathematical characteristics, for example conservation properties, that may be exploited in their numerical integration. In this thesis we investigate whether we can also exploit these conservation properties within a data assimilation scheme by adding an additional constraint that imposes these conservation properties. In order to consider this we construct variational data assimilation schemes for two Hamiltonian systems, the two and three-body problems.

1.2 Thesis Summary

We begin in Chapter 2 by introducing the mathematics of Hamiltonian systems and discussing the inherent conservation properties of such problems. We will then discuss the numerical methods that can be implemented when modelling Hamiltonian problems, with specific mention of symplectic methods. We end this chapter by demonstrating that the atmosphere can be described as a Hamiltonian system.

In Chapter 3 we will briefly discuss the development of data assimilation and describe in detail the mathematics of four dimensional variational (4D Var) data assimilation, the technique we will be using in this investigation. We also include a discussion of the stages required in the construction of the 4D Var scheme and the tests that are required to validate the algorithm. We will then introduce the addition of weak constraints to the system and discuss current research using this formulation of the 4d Var problem.

Following on from this in Chapter 4 we describe the development, verification and investigation of the 4d Var scheme for the two-body problem. We will begin by deriving both the continuous and discrete systems. We will then test the model, with particular attention to the conservation properties of the discrete system. We also detail the stages required in the construction of the 4d Var algorithm, including the development and testing of the linear model. We will finally discuss the experiments we have carried out for this two-body data assimilation scheme. This includes considering the effect of our choice of observations and the addition of weak constraints. These are imposed at the initial time only and include constraints that make use of the conservation properties of the system.

In Chapter 5 we introduce the second of our Hamiltonian models. We will discuss both the continuous and discrete equations for the three-body problem before carrying out a series of experiments with the numerical model in order to find the most appropriate configuration of the three-body problem. We will then consider the sensitivity and stability of the resulting system. We provide details relating to the construction of the data assimilation scheme and repeat the validation tests we carried out for the two-body case. We will describe the experiments we carried out using this scheme, again considering the effect of having an incomplete set of observations. We also consider implementing weak constraints at the initial time, again using the conservation property of the system. We investigate whether these additional constraints can improve the solution where observations alone have failed.

Finally in Chapter 6 we summarise the work we have done and discuss any conclusions that can be drawn from the work. We also give suggestions for further work on this problem.

Chapter 2

Hamiltonian Systems

Recent developments in atmospheric dynamics have led to the construction of a Hamiltonian formulation of the equations governing the evolution of the atmosphere [4]. We review this in more detail in section 2.3 of this chapter. This work provides the motivation for this thesis in which we consider whether the intrinsic characteristics of such a Hamiltonian system are important when creating a data assimilation scheme. In this chapter we discuss the mathematics of a general Hamiltonian system and review the numerical methods that are suitable for constructing models for Hamiltonian problems.

2.1 Mathematical Formulation of Hamiltonian Systems

In this section we provide the derivation of the Hamiltonian equations of motion. These are a set of coupled equations that govern the position, $\boldsymbol{q}$, and momentum, $\boldsymbol{p}$, of a particle or body using a generalised co-ordinate system. This differs from the Newtonian formulation which uses Cartesian coordinates, $\boldsymbol{x}$, and their time derivatives, $\dot{\boldsymbol{x}}$, in the equations of motion.

We consider an example for which two important assumptions have been made; the system is *conservative* and the co-ordinate transformation is time independent, i.e. the axes are fixed. Although we do not derive the most general formulation of Hamiltonian mechanics, this description is sufficient for the Hamiltonian systems considered here. Details of the derivation of the general problem can be found in [19, 25].

We begin by considering Newton's equations of motion for a general system of N particles [26],

$$\boldsymbol{F}_i = \frac{\mathrm{d}}{\mathrm{d}t}(m_i \dot{\boldsymbol{x}}_i) \qquad \boldsymbol{F}_i, \dot{\boldsymbol{x}}_i \in \Re^d \quad (i = 1, 2, \ldots N) \tag{2.1}$$

where $\boldsymbol{F}_i$ denotes the vector of forces acting on the ith particle, m_i is the mass of the ith particle and $\dot{\boldsymbol{x}}_i$ is the time derivative of the position vector for the ith particle, $\boldsymbol{x}_i$. Each vector is of length d, the space dimension of the problem.

For this system the kinetic energy, T, is defined as

$$T = \sum_i \frac{1}{2} m_i \dot{\boldsymbol{x}}_i \cdot \dot{\boldsymbol{x}}_i \qquad \dot{\boldsymbol{x}}_i \in \Re^d \quad (i = 1, 2, \ldots N). \tag{2.2}$$

From equations (2.1) and (2.2) we can see that

$$\boldsymbol{F}_i = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{\boldsymbol{x}}_i}\right) \qquad \boldsymbol{F}_i, \dot{\boldsymbol{x}}_i \in \Re^d \quad (i = 1, 2, \ldots N), \tag{2.3}$$

where the system has potential energy $V = V(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N)$. In this thesis we consider conservative systems, where by definition [26] the forces, $\boldsymbol{F}_i$, are given by

$$\boldsymbol{F}_i = -\frac{\partial V}{\partial \boldsymbol{x}_i} \qquad \boldsymbol{F}_i, \boldsymbol{x}_i \in \Re^d \quad (i = 1, 2, \ldots N). \tag{2.4}$$

We thus have the following relationship between the kinetic energy and potential energy of the particles,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{\boldsymbol{x}}_i}\right) = -\frac{\partial V}{\partial \boldsymbol{x}_i} \qquad \boldsymbol{x}_i, \dot{\boldsymbol{x}}_i \in \Re^d \quad (i = 1, 2, \ldots N). \tag{2.5}$$

We now introduce the generalised coordinates $\boldsymbol{q}$. If we were to consider the most general form then position, $\boldsymbol{q}$, would depend explicitly on time, thus allowing for a moving co-ordinate system. However, we are considering a transformation from one stationary coordinate system to another, therefore $\boldsymbol{q} \equiv \boldsymbol{q}(\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \dots \boldsymbol{x}_N^T)^T$, where $\boldsymbol{q} \in \Re^{Nd}$.

As demonstrated in [26], we can reformulate (2.3) in this generalised coordinate system. We thus have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{\boldsymbol{q}}}\right) = \frac{\partial T}{\partial \boldsymbol{q}} - \frac{\partial V}{\partial \boldsymbol{q}}, \tag{2.6}$$

where $\dot{\boldsymbol{q}}$ is the time derivative of $\boldsymbol{q}$. Such conservative systems, i.e. those with no external forcings, can be described using the Lagrangian function, $\mathcal{L} = \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})$, defined as

$$\mathcal{L} = T - V \tag{2.7}$$

where $T = T(\boldsymbol{q}, \dot{\boldsymbol{q}})$ and $V = V(\boldsymbol{q})$. Therefore equation (2.6) is now

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}\right) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}}. \tag{2.8}$$

We can arrive at the same result by starting from *Hamilton's Principle* and using a calculus of variations approach to investigate the dynamics. Hamilton's Principle states that:

The motion of the system from time t_1 to time t_2 is such that the line integral

$$I = \int_{t_1}^{t_2} \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})\,\mathrm{d}t \tag{2.9}$$

is an extremum following the path of motion [19].

We now consider the first variation of I for fixed t_1 and t_2, with boundary conditions imposed such that $\delta\boldsymbol{q} = 0$ and $\delta\dot{\boldsymbol{q}} = 0$ at t_1 and t_2. If Hamilton's

Principle holds then δI is zero [19] and we have

$$\delta I = \delta \int_{t_1}^{t_2} \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \, \mathrm{d}t = \int_{t_1}^{t_2} \delta \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}) \, \mathrm{d}t = 0. \tag{2.10}$$

From [15] we have

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}} \cdot \delta \boldsymbol{q} + \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}} \cdot \delta \dot{\boldsymbol{q}}, \tag{2.11}$$

and thus (2.10) becomes

$$\delta I = \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{q}} \cdot \delta \boldsymbol{q} + \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}} \cdot \delta \dot{\boldsymbol{q}} \right) \mathrm{d}t = 0. \tag{2.12}$$

But $\delta \dot{\boldsymbol{q}} = \frac{\mathrm{d}}{\mathrm{d}t}(\delta \boldsymbol{q})$, and integrating by parts gives us,

$$\delta I = \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{q}} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}} \right) \right) \cdot \delta \boldsymbol{q} \, \mathrm{d}t = 0, \tag{2.13}$$

since the variation is zero at the boundaries. As the variation in $\boldsymbol{q}$ is arbitrary it must follow that,

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}} \right) - \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}} = 0. \tag{2.14}$$

These are the Euler-Lagrange equations - second-order equations in time describing the evolution of the system.

We arrive at the *Hamiltonian* formulation of (2.14) by defining generalised momenta,

$$\boldsymbol{p} = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}. \tag{2.15}$$

We now use this to relate the concept of a Hamiltonian to the Lagrangian. We consider,

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}} \cdot \frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} + \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}} \cdot \frac{\mathrm{d}\dot{\boldsymbol{q}}}{\mathrm{d}t}. \tag{2.16}$$

Using (2.14) and rearranging, this can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}} \cdot \dot{\boldsymbol{q}} - \mathcal{L} \right) = 0. \tag{2.17}$$

Using the definition for generalised momenta from (2.15) we thus have,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{p}\cdot\dot{\boldsymbol{q}}-\mathcal{L}\right)=0. \tag{2.18}$$

Since $\mathcal{L}$, $\boldsymbol{q}$ and $\boldsymbol{p}$ do not depend explicitly on time, t, then it follows that

$$\boldsymbol{p}\cdot\dot{\boldsymbol{q}}-\mathcal{L}=\text{constant}. \tag{2.19}$$

We can therefore define

$$H(\boldsymbol{p},\boldsymbol{q})=\boldsymbol{p}\cdot\dot{\boldsymbol{q}}-\mathcal{L}(\boldsymbol{q},\dot{\boldsymbol{q}}). \tag{2.20}$$

This is known as the *Hamiltonian* and is a constant following the evolution of the system.

We can use (2.20) to derive a set of first order equations which are equivalent to the second order Euler-Lagrange equations (2.14) [26]. From (2.20) it follows that

$$\mathrm{d}H=\dot{\boldsymbol{q}}\cdot\mathrm{d}\boldsymbol{p}+\boldsymbol{p}\cdot\mathrm{d}\dot{\boldsymbol{q}}-\mathrm{d}\mathcal{L}. \tag{2.21}$$

Using the definition of the generalised momenta given in equation (2.15), (2.21) becomes,

$$\mathrm{d}H=\dot{\boldsymbol{q}}\cdot\mathrm{d}\boldsymbol{p}-\frac{\partial\mathcal{L}}{\partial\boldsymbol{q}}\cdot\mathrm{d}\boldsymbol{q}. \tag{2.22}$$

Substituting (2.15) into the Euler-Lagrange equation (2.14) gives $\dot{\boldsymbol{p}}=\frac{\partial\mathcal{L}}{\partial\boldsymbol{q}}$. If we substitute this into (2.22) we have,

$$\mathrm{d}H=-\dot{\boldsymbol{p}}\cdot\mathrm{d}\boldsymbol{q}+\dot{\boldsymbol{q}}\cdot\mathrm{d}\boldsymbol{p}. \tag{2.23}$$

In addition we also have that

$$\mathrm{d}H=\frac{\partial H}{\partial\boldsymbol{q}}\cdot\mathrm{d}\boldsymbol{q}+\frac{\partial H}{\partial\boldsymbol{p}}\cdot\mathrm{d}\boldsymbol{p}. \tag{2.24}$$

Comparison of the coefficients of equations (2.23) and (2.24) gives the following equations,

$$\dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}}, \tag{2.25}$$

$$\dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}}. \tag{2.26}$$

This set of first order equations is sufficient to describe the evolution of the system, and is known as the *canonical form* for a Hamiltonian system. The Hamiltonian formulation of a problem can prove to be a useful description since the equations are first order and it is independent of the coordinate system used.

As illustrated in [6], by defining vector $\boldsymbol{u} = (\boldsymbol{q}^T, \boldsymbol{p}^T)^T$ where $\boldsymbol{q}, \boldsymbol{p} \in \mathbb{R}^{Nd}$, we can rewrite these equations as

$$\dot{\boldsymbol{u}} = \boldsymbol{J}^{-1} \nabla_{\boldsymbol{u}} H, \tag{2.27}$$

where $\nabla_{\boldsymbol{u}} = (\frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_d}, \frac{\partial}{\partial q_1}, \ldots, \frac{\partial}{\partial q_d})$ and J is the so-called *symplectic operator* [34] defined by

$$\boldsymbol{J} = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix}. \tag{2.28}$$

We can say that the system

$$\dot{\boldsymbol{u}} = \boldsymbol{f}(\boldsymbol{u}) \tag{2.29}$$

is Hamiltonian if

$$\boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{J}^{-1} \nabla_{\boldsymbol{u}} H. \tag{2.30}$$

For the conservative system we have considered, $\boldsymbol{\mathcal{L}} = T(\dot{\boldsymbol{q}}) - V(\boldsymbol{q})$, thus (2.20) becomes

$$H = \boldsymbol{p} \cdot \dot{\boldsymbol{q}} - (T - V). \tag{2.31}$$

Using the definition of $\boldsymbol{p}$ from (2.15) we have

$$\boldsymbol{p} = \frac{\partial T}{\partial \dot{\boldsymbol{q}}}, \tag{2.32}$$

since $V = V(\boldsymbol{q})$ and therefore does not depend on $\dot{\boldsymbol{q}}$. Using this definition of the momentum with (2.31) we find that the Hamiltonian, H, is given by

$$H = T + V. \tag{2.33}$$

Thus where we have a conservative system with a generalised co-ordinate system that is fixed, the Hamiltonian is the total energy, and this property is conserved following the motion.

2.2 Numerical Methods for Hamiltonian Problems

When considering which numerical scheme to use to model a given problem, we are usually concerned with finding one that is not only stable, but also minimises the local truncation error arising from the discretisation, i.e. one that is more accurate. However such methods, although locally accurate, do not always account for the global features of the system. Consider the problem of two bodies in mutual orbit. As we will see in Chapter 4, intrinsic to this problem is the fact that *energy is conserved*. It is therefore essential that the numerical model captures this property. This can be illustrated by considering the effect of energy loss in a numerical model of the solar system [6]. The orbiting bodies would spiral inwards, which is physically incorrect - planets do not spiral into the sun. Hence for such a system local accuracy is less important than capturing the conservation properties.

2.2.1 Geometric Methods

Geometric methods have been developed specifically to preserve the qualitative features of a system. These may include geometrical structure, conservation laws, symmetries and asymptotic behaviour. Considering these properties when designing numerical schemes has proved successful in many fields including astronomy and molecular dynamics [27, 44]. For Hamiltonian problems *symplectic* methods have proved particularly effective at preserving the conservation properties of the system.

As in Section 2.1 we define the vector $\boldsymbol{u}$ such that $\boldsymbol{u} = (\boldsymbol{q}^T, \boldsymbol{p}^T)^T$ and $\boldsymbol{q}, \boldsymbol{p} \in \mathbb{R}^{Nd}$, $\boldsymbol{q}$ and $\boldsymbol{p}$ are position and momentum as before. The evolution is given by

$$\dot{\boldsymbol{u}} = \boldsymbol{f}(\boldsymbol{u}), \tag{2.34}$$

where $\boldsymbol{f}(\boldsymbol{u})$ satisfies (2.30).

An additional feature of a Hamiltonian system is that the flow is *symplectic*. We can consider that the solution of (2.34) at a given time, T, induces a transformation, ϕ, of $\boldsymbol{u}$ from the initial time, $t = 0$ to $t = T$. This mapping is symplectic if

$$\phi'^T J \phi' = J, \tag{2.35}$$

where ϕ' is the Jacobian of ϕ, and J is the skew-symmetric matrix defined in (2.28). The proof for this is given in [6]. We will see that the symplectic nature of the flow of a Hamiltonian system will prove useful in the development of numerical schemes.

We now consider a numerical approximation for the system given in (2.34). We consider a discrete scheme that induces a mapping Φ_h of our discrete vector $\boldsymbol{U}$, where $\boldsymbol{U} = (\boldsymbol{Q}^T, \boldsymbol{P}^T)^T$, from the initial to the final time. This is symplectic if it satisfies (2.35). Such schemes are known as symplectic numerical methods and prove to be excellent for modelling Hamiltonian systems as they preserve many of the qualitative features. In fact we observe that

> “a symplectic discretisation of a Hamiltonian problem is a Hamiltonian perturbation of the original” [6]

We can thus say that whereas non-geometric methods produce an approximation to the solution, symplectic methods *exactly* solve an approximate Hamiltonian system.

2.2.2 The Störmer-Verlet Method

The Störmer-Verlet is a second-order accurate, symplectic scheme. This method makes use of Lemma 3.1 in [6], which states that if maps ω and ψ are symplectic, then the composition of the two, $\omega \circ \psi$ is also symplectic [6]. Thus we can decompose our discrete map Φ_h into a composition of simpler, symplectic flows,

$$\Phi_h = \Phi_{1,h} \circ \Phi_{2,h} \circ \Phi_{3,h} \ldots . \tag{2.36}$$

Each of these simpler flows can be chosen to reflect a simpler integration of the original problem. We can use this property to define various *splittings* of the problem. For a more complete discussion of splitting methods see [6].

We consider a separable Hamiltonian problem such that $H(\boldsymbol{p}, \boldsymbol{q}) = H_1(\boldsymbol{p}) + H_2(\boldsymbol{q})$. For such systems the continous equations become,

$$\dot{\boldsymbol{q}} = \frac{\partial H_1(\boldsymbol{p})}{\partial \boldsymbol{p}} \tag{2.37}$$

$$\dot{\boldsymbol{p}} = -\frac{\partial H_2(\boldsymbol{q})}{\partial \boldsymbol{q}} \tag{2.38}$$

Applying a *Strang* splitting [6] to this system results in the Störmer-Verlet method [20] given by the following equations,

$$\begin{aligned}
\boldsymbol{P}^{n+\frac{1}{2}} &= \boldsymbol{P}^n - \frac{h}{2}\frac{\partial H_2(\boldsymbol{Q}^n)}{\partial \boldsymbol{Q}^n} && (2.39)\\
\boldsymbol{Q}^{n+1} &= \boldsymbol{Q}^n + h\frac{\partial H_1(\boldsymbol{P}^{n+\frac{1}{2}})}{\partial \boldsymbol{P}^{n+\frac{1}{2}}} && (2.40)\\
\boldsymbol{P}^{n+1} &= \boldsymbol{P}^{n+\frac{1}{2}} - \frac{h}{2}\frac{\partial H_2(\boldsymbol{Q}^{n+1})}{\partial \boldsymbol{Q}^{n+1}}, && (2.41)
\end{aligned}$$

where the use of upper case letters indicates a discrete system. $\boldsymbol{P}^n$ is the momentum at the nth timestep, $\boldsymbol{Q}^n$ the position, h the length of the timestep.

We can see that this splitting is symplectic by considering each of these discrete equations. In (2.39) we change momentum $\boldsymbol{P}$ only, and $\boldsymbol{Q}$ remains constant. The change in $\boldsymbol{P}$ is given by (2.38). We can write this as a Hamiltonian system in the following way

$$\begin{aligned}
\dot{\boldsymbol{p}} &= -\frac{\partial H_2(\boldsymbol{q})}{\partial \boldsymbol{q}} && (2.42)\\
\dot{\boldsymbol{q}} &= \frac{\partial H_2(\boldsymbol{p})}{\partial \boldsymbol{p}}. && (2.43)
\end{aligned}$$

Since H_2 is independent of $\boldsymbol{p}$ then $\frac{\partial H_2(\boldsymbol{p})}{\partial \boldsymbol{p}} = 0$. Thus, as the discrete equations illustrate, the position is not changed by the Hamiltonian system given by (2.42) and (2.43).

Similarly, we consider the step produced by (2.40). Here we change position, $\boldsymbol{Q}$, and the momentum remains constant. This is given by (2.37. Again this can be written as a Hamiltonian system

$$\begin{aligned}
\dot{\boldsymbol{q}} &= \frac{\partial H_2(\boldsymbol{p})}{\partial \boldsymbol{p}} && (2.44)\\
\dot{\boldsymbol{p}} &= -\frac{\partial H_2(\boldsymbol{q})}{\partial \boldsymbol{q}}. && (2.45)
\end{aligned}$$

As H_1 is independent of $\boldsymbol{q}$ then $\frac{\partial H_2(\boldsymbol{q})}{\partial \boldsymbol{q}} = 0$ and the momentum remains constant.

We can follow a similar argument for the final discrete equation. We can thus see that each of the discrete equations forms a Hamiltonian system, and thus the flow for each is symplectic. Thus the mapping for each of the variables, from one timestep to the next, is symplectic. Using the result given in (2.36), we can thus conclude that the method is symplectic.

2.3 The Atmosphere as a Hamiltonian System

Hamiltonian methods are tried and tested in the field of classical mechanics, however their extension to fluid mechanics is a relatively recent development [43]. We begin by considering the Lagrangian description of a fluid. Here each fluid particle is labelled $\boldsymbol{a} = (a, b, c)$ which can be defined by the initial positions. This label is fixed and travels with the flow of the fluid. We introduce a further independent variable τ that denotes time such that (a, b, c) remain fixed by $\frac{\partial}{\partial \tau}$.

By convention, the canonical variables for Hamiltonian fluid dynamics are given by $\boldsymbol{x}$ and $\boldsymbol{u}$ for position and momentum respectively. We recall that the Lagrangian for a system of N discrete particles is given by,

$$\mathcal{L} = T - V, \tag{2.46}$$

where the kinetic energy, T, is defined by (2.2), and V is the potential energy. By allowing N to increase without limit and the distances between particles to become arbitrarily small we can find a Lagrangian to represent a continuous fluid [34]. The kinetic energy becomes

$$T = \int\int\int \left(\frac{1}{2} \frac{\partial \boldsymbol{x}}{\partial \tau} . \frac{\partial \boldsymbol{x}}{\partial \tau} \right) \mathrm{d}a \mathrm{d}b \mathrm{d}c \tag{2.47}$$

and the potential energy is

$$V = \int\int\int \left(E(\alpha, S) + \Phi(\boldsymbol{x})\right) \mathrm{d}a\mathrm{d}b\mathrm{d}c, \tag{2.48}$$

where E is the internal energy and Φ represents the external energy from, say, gravity or other external forces. The internal energy is a function of the specific volume, $\alpha = \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{a}}$, and S is the specific entropy.

As we have seen in Section 2.1 Hamilton's principle requires that $\delta I = 0$ where

$$I = \int \mathcal{L} dt. \tag{2.49}$$

Following a similar argument we obtain the Lagrangian equations for a fluid

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{x}}} - \frac{\partial \mathcal{L}}{\partial \boldsymbol{x}} = 0, \tag{2.50}$$

where

$$\mathcal{L} = \int\int\int \left(\frac{1}{2}\frac{\partial \boldsymbol{x}}{\partial \tau}.\frac{\partial \boldsymbol{x}}{\partial \tau} - \left(E(\alpha, S) + \Phi(\boldsymbol{x})\right)\right) \mathrm{d}a\mathrm{d}b\mathrm{d}c. \tag{2.51}$$

We introduce generalised momentum

$$\boldsymbol{u} = \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{x}}} \tag{2.52}$$

By analogy with (2.20) the Hamiltonian is defined by

$$H = \int\int\int \boldsymbol{u} \cdot \dot{\boldsymbol{x}} \mathrm{d}\boldsymbol{a} - \mathcal{L}, \tag{2.53}$$

and the canonical equations are given by (2.25) and (2.26).

2.3.1 Example: Shallow Water Equations

We have demonstrated that Hamiltonian methods can be extended to fluids. We now show that they can be used for equations that govern the evolution of the atmosphere. One such model is given by the shallow water equations. These describe the motion of a shallow layer of incompressible fluid, in two dimensions, in a rotating coordinate system with a flat topology. The height of the free surface is h. The equations of motion for the shallow water system are [34]

$$\frac{\partial^2 x}{\partial \tau^2} - f\frac{\partial y}{\partial \tau} + g\frac{\partial h}{\partial x} = 0 \tag{2.54}$$

$$\frac{\partial^2 y}{\partial \tau^2} + f\frac{\partial x}{\partial \tau} + g\frac{\partial h}{\partial y} = 0, \tag{2.55}$$

where g is the gravitational potential and f is the planetary vorticity.

For a rotating frame we introduce $\boldsymbol{W} = \boldsymbol{\Omega}\times\boldsymbol{x}$ - the velocity of an atmosphere rotating with the Earth - where $\boldsymbol{\Omega}$ is the angular velocity of the rotating frame. The Lagrangian is thus [34]

$$\mathcal{L} = \int\int \frac{1}{2}\left(\frac{\partial \boldsymbol{x}}{\partial \tau}.\frac{\partial \boldsymbol{x}}{\partial \tau} + 2\boldsymbol{W}.\frac{\partial \boldsymbol{x}}{\partial \tau} + \boldsymbol{W}.\boldsymbol{W} - gh\right) \mathrm{d}a\mathrm{d}b, \tag{2.56}$$

since we are considering a two dimensional system.

Using (2.15) we derive our generalised momentum,

$$\boldsymbol{u} = \dot{\boldsymbol{x}} + \boldsymbol{W}. \tag{2.57}$$

Thus, using the definition of the Hamiltonian from (2.53), we find

$$H = \int\int \left(\frac{1}{2}\boldsymbol{u}.\boldsymbol{u} - \boldsymbol{u}.\boldsymbol{W} + \frac{1}{2}gh\right) \mathrm{d}a\mathrm{d}b. \tag{2.58}$$

Therefore, using (2.44) and (2.45), we can derive the canonical equations for (2.54) and (2.55)

$$\dot{\boldsymbol{x}} = \boldsymbol{u} - \boldsymbol{W} \tag{2.59}$$

$$\dot{\boldsymbol{u}} = -(\boldsymbol{\Omega}\times\boldsymbol{u} + g\nabla h). \tag{2.60}$$

These, along with the continuity equation

$$\frac{\partial h}{\partial \tau} + h\nabla\cdot\boldsymbol{v} = 0 \tag{2.61}$$

where $\boldsymbol{v} = (\frac{\partial x}{\partial \tau}, \frac{\partial y}{\partial \tau}$, form the shallow water system.

This example demonstrates that it is possible to rewrite the equations of motion for a model of the atmosphere in Hamiltonian form. It can also be shown [37] that Hamiltonian balanced models can be constructed from more complicated unbalanced models. As we have discussed in the previous section certain geometric methods are very successful for such models. Therefore if we can construct a Hamiltonian model for a given problem we can take advantage not only of these numerical methods but also of the inherent conservation and symmetry characteristics of Hamiltonian systems.

2.4 Summary

In Chapter 2 we have considered the mathematics of Hamiltonian systems. By beginning with Newton's equations of motion we have shown that a first order, Hamiltonian system can be derived. This system has notable geometrical features and conserves the Hamiltonian following the evolution of the system. For conservative problems in a fixed co-ordinate system, this Hamiltonian is the total energy of the system.

We have discussed the choice of numerical method that is appropriate for such problems. We have seen that symplectic schemes, which are a type of geometric method, are conserving models and as such are a good choice for Hamiltonian problems.

Finally we have provided the motivation for this research by demonstrating that equations that model the evolution of the atmosphere can be written in Hamiltonian form.

In this thesis we use a simple Hamiltonian system and consider whether the characteristics of this problem can be exploited within the data assimilation algorithm. Therefore, in the following chapter, we consider the theory and practical details of data assimilation with specific reference to the four-dimensional variational approach. This is necessary in order that we can construct a data assimilation system for our Hamiltonian models later in the thesis.

Chapter 3

Data Assimilation

In this thesis we consider whether conservation properties of a given system can be exploited when constructing a *data assimilation algorithm.* Data assimilation is a method that provides a best estimate of the state of a system by incorporating observational data into a numerical model. Several different data assimilation techniques have been developed to solve this problem, varying in computational cost, optimality and speed.

The aim of this thesis is to investigate data assimilation techniques for a simple system with the intention that the results may have implications for real world meteorological problems. We therefore consider the methods commonly used in this field of study. In this chapter we introduce the basic ideas of data assimilation and briefly discuss the various methods that can be used. We then give a more detailed account of the assimilation technique we have chosen to use for this research - four-dimensional variational data assimilation (4d Var).

3.1 An Introduction to Data Assimilation

To achieve the best possible estimate of the state of a system we can use data assimilation methods to make use of the information that we have. This information includes any observations that are available, equations governing the dynamics of the system, information from previous forecasts, and statistical information relating to errors in the observations and previous forecasts. In addition we may wish to exploit any physical relationships between variables and physical constraints on the evolution of the system.

For atmospheric models we have many observations available. However they are insufficient in number to fully describe its state and this data can prove difficult to use. For example satellite measurements give the integrated sum of the radiation emitted over the entire vertical profile whereas the numerical model includes many vertical levels. In addition, observations are often of different variables to those that we want to use in the model. A further obstacle with all observations is the error associated with these measurements. A good assimilation scheme will attempt to include these complexities [13].

For most systems we have a set of differential equations that govern the evolution of the state of the system. A data assimilation scheme will integrate the observations into an appropriate numerical model for these equations. With some data assimilation techniques it is also possible to add a physical constraint, for example it might be advantageous to constrain the motion to lie on, or close to, a particular surface in phase space.

A further feature of data assimilation schemes is that we can make use of previous information. This is known as the *background*, and in most cases this is the forecast given by an earlier cycle. The addition of this term means that information from observations assimilated in the past is not lost. Again there are errors associated with this that must be considered.

As we have discussed we can obtain a large amount of detail relating to a system, but it is important that this is used appropriately. A very simple data assimilation scheme would directly replace the value predicted by the model equations with the observation at that point. This, however, may be inconsistent with the model values at neighbouring points. Better data assimilation schemes provide a solution that is consistent with both the model predictions and the observations.

A successful data assimilation scheme should include the following [5]

1. A good background field, typically a forecast from a previous data assimialtion cycle.
2. A method that provides a compromise between the background and the observations. We want the scheme to produce a solution that is close to the data that is the most reliable.
3. A smooth solution as we know that the real field is smooth.
4. A solution that respects known physical properties and dynamics.

In addition we also want to include any errors associated with the observations and the background field.

3.1.1 Basic Concepts

We now define several terms that are frequently used when describing data assimilation methods and their solutions [5].

The *analysis* is the best estimate of the state of the system at a given time.

The *state vector*, $\boldsymbol{x}$, describes the column vector of variables that represent the state of the system and defines our model space. In addition we define $\boldsymbol{x}_t$ to be the true state at the analysis time, $\boldsymbol{x}_b$ the background estimate of the true state before the assimilation is carried out, and $\boldsymbol{x}_a$ is the analysis.

We also have a vector of *observations*, $\boldsymbol{y}$, which defines our observation space. In general we have fewer observations than model variables and as we have already stated the variables are often different. Thus we cannot compare the two directly and we must use an observation operator, H. This operator is applied to the state vector to produce model equivalents for the observations.

The vector of *departures* measures the difference between the observations and the state vector, $\boldsymbol{y} - H(\boldsymbol{x})$. This is sometimes referred to as the innovation vector.

In addition we have error covariance matrices $\mathbf{B}$ and $\mathbf{R}$, describing the statistics relating to the errors of the background state and the observations respectively.

In general we can classify different data assimilation techniques as sequential or non-sequential, intermittent or continuous. Sequential assimilation considers only observations taken before the time of the analysis whereas non-sequential techniques include observations taken after the analysis time. Intermittent methods treat all observations as if they occurred at the same time and assimilate them at that point. In contrast continuous techniques assimilate the observation at the time that it was measured.

3.1.2 Variational Data Assimilation

Current operational numerical weather prediction methods have generally moved towards *variational* tecniques. Such methods involve the definition of a *cost func-*

tion, J, that is some measure of the distance between the observations and the model state and between the background and the model state. This is then minimised in order to find a solution that lies close to both - this solution is the *analysis*.

One such method is three dimensional variational data assimilation (3d Var) which was used operationally at the Met Office UK between 1999 and 2004 [32]. This uses observations that are distributed in space, but are assumed to be valid at the same point in time. This approximation is computationally less expensive than other techniques. However there is a loss of some information provided by the observations.

An extension of this is four-dimensional data assimilation (4d Var), where observations are distributed in both space and time [5]. The addition of the time dimension treats the observations more realistically. A variation of this technique is currently operational at both the European Centre for Medium Range Weather Forecasting (ECMWF) and the UK Met Office. In both cases it is the incremental form of 4d Var that is employed. Details of this approach to the 4d Var method can be found in [10]. In this thesis we will we use the *full* 4d Var method, which we now consider in more detail.

3.2 The Mathematical Formulation of 4d Var

The mathematics of the 4d Var problem can be considered using linear algebra, or by using the technique of calculus of variations. In most practical applications, the linear algebra approach is used. However it can also be useful to look at the variational formulation.

3.2.1 The Continuous Problem

In order to illustrate how the cost function, model equations and adjoint equations are related, we initially consider the continuous problem.

We have a general system with a cost functional given by

$$J = \int_{t_0}^{t_1} F(t, \boldsymbol{x}, \dot{\boldsymbol{x}}). \tag{3.1}$$

We wish to minimise J subject to the constraint that

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}), \tag{3.2}$$

where $\cdot \equiv \frac{d}{dt}$, t is an independent variable and $\boldsymbol{x}$ is dependent on t. The value of J depends on the path between the two end points.

For 4d Var, $F(t, \boldsymbol{x}, \dot{\boldsymbol{x}})$ is a function that measures the difference between the observations and the model state and between the background and the model state. The constraint is given by the equations of motion for the system under consideration.

We can rewrite this problem by defining the *Lagrangian* functional, $\boldsymbol{\mathcal{L}}$, which takes into account the additional information provided by the constraint.

$$\boldsymbol{\mathcal{L}} = \int_{t_0}^{t_1} \left(F(t, \boldsymbol{x}, \dot{\boldsymbol{x}}) + \boldsymbol{\lambda}(t)(\dot{\boldsymbol{x}} - f(\boldsymbol{x})) \right) dt, \tag{3.3}$$

where $\boldsymbol{\lambda}(t)$ is a vector of Lagrange multipliers. We set

$$G(\boldsymbol{x}, \boldsymbol{\lambda}, \dot{\boldsymbol{x}}, \dot{\boldsymbol{\lambda}}) = F(t, \boldsymbol{x}, \dot{\boldsymbol{x}}) + \boldsymbol{\lambda}(t)(\dot{\boldsymbol{x}} - f(\boldsymbol{x})). \tag{3.4}$$

Thus our Lagrangian becomes

$$\boldsymbol{\mathcal{L}} = \int_{t_0}^{t_1} G(\boldsymbol{x}, \boldsymbol{\lambda}, \dot{\boldsymbol{x}}, \dot{\boldsymbol{\lambda}}) dt. \tag{3.5}$$

It is *this* functional that we now want to minimise as it is not only a measure of the original function, but it also imposes the constraint required by the original problem.

To find the minimum we require the first variation of the functional to be zero, i.e.

$$\delta\mathcal{L} = \delta \int_{t_0}^{t_1} G(\boldsymbol{x}, \boldsymbol{\lambda}, \dot{\boldsymbol{x}}, \dot{\boldsymbol{\lambda}}) = \int_{t_0}^{t_1} \delta G(\boldsymbol{x}, \boldsymbol{\lambda}, \dot{\boldsymbol{x}}, \dot{\boldsymbol{\lambda}}) = 0. \tag{3.6}$$

From (3.6) we have

$$\delta\mathcal{L} = \int_0^t (\delta\boldsymbol{x} G_{\boldsymbol{x}} + \delta\dot{\boldsymbol{x}} G_{\dot{\boldsymbol{x}}} + \delta\boldsymbol{\lambda} G_\lambda + \delta\dot{\boldsymbol{\lambda}} G_{\dot{\lambda}})dt = 0, \tag{3.7}$$

where $G_{\boldsymbol{x}}$ is the differential of G with respect to $\boldsymbol{x}$, $G_{\dot{\boldsymbol{x}}}$ is the diffferential with respect to $\dot{\boldsymbol{x}}$ and G_λ and $G_{\dot{\lambda}}$ are similarly defined. Integrating by parts and applying boundary conditions, leads to

$$\delta\mathcal{L} = \int_0^t \left(\left(G_{\boldsymbol{x}} - \frac{d}{dt} G_{\dot{\boldsymbol{x}}} \right) \delta\boldsymbol{x} + \left(G_\lambda + \frac{d}{dt} G_{\dot{\lambda}} \right) \delta\boldsymbol{\lambda} \right) dt = 0. \tag{3.8}$$

Since $\boldsymbol{x}$ and $\boldsymbol{\lambda}$ are independent, we can thus deduce the following equations that are necessary to ensure $\delta\mathcal{L} = 0$,

$$G_{\boldsymbol{x}} - \frac{d}{dt} G_{\dot{\boldsymbol{x}}} = 0 \tag{3.9}$$

$$G_\lambda - \frac{d}{dt} G_{\dot{\lambda}} = 0. \tag{3.10}$$

These are known as the Euler-Lagrange equations. The second of these gives the continuous model equations describing the evolution of the system. The first gives a set of equations that are referred to as the *adjoint* equations. As we will see in the following section, the discrete equivalent of these are required to minimise the discrete cost function. Here, satisfying both (3.9) and (3.10) will result in finding the minimum of the Lagrangian, and therefore the minimum of the original objective function, J, subject to the original constraint that $\dot{\boldsymbol{x}} = f(\boldsymbol{x})$.

3.2.2 The Discrete Problem

We now consider the discrete 4d Var problem.

We want to minimise the cost functional

$$J(\boldsymbol{x}_0) = (\boldsymbol{x}_0 - \boldsymbol{x}_b)^T \boldsymbol{B}^{-1} (\boldsymbol{x}_0 - \boldsymbol{x}_b) + \sum_{n=0}^{N} (\boldsymbol{y}_n - H_n[\boldsymbol{x}_n])^T \boldsymbol{R}_n^{-1} (\boldsymbol{y}_n - H_n[\boldsymbol{x}_n]), \quad (3.11)$$

subject to the strong constraint that

$$\boldsymbol{x}_n = f(\boldsymbol{x}_{n-1}), \quad (3.12)$$

where the subscript n denotes quantities at observation time n from $n = 0$ to $n = N$ and f is a nonlinear function. Here (3.12) describes the dynamics of the discrete system.

Here it is assumed that $f(\boldsymbol{x}_n)$ is a perfect representation of the equations of motion given by (3.12). This assumption is also made in current operational systems. The inclusion of an additional term in the cost function to account for model error is being considered [39].

To understand how we minimise the discrete cost function, we can follow a similar argument as for the continuous problem. We define a discrete Lagrangian which includes the additional information provided by the constraint.

$$\mathcal{L}_n = \sum_{n=0}^{N} \left(F_n(\boldsymbol{x}_n, \dot{\boldsymbol{x}}_n) + \boldsymbol{\lambda}_n (\dot{\boldsymbol{x}}_n - f_n(\boldsymbol{x}_n)) \right), \quad (3.13)$$

where $\boldsymbol{\lambda}_n$ is the vector of the discrete adjoint variables. We set

$$G_n(\boldsymbol{x}_n, \boldsymbol{\lambda}_n, \dot{\boldsymbol{x}}_n, \dot{\boldsymbol{\lambda}}_n) = F_n(\boldsymbol{x}_n, \dot{\boldsymbol{x}}_n) + \boldsymbol{\lambda}_n (\dot{\boldsymbol{x}}_n - f_n(\boldsymbol{x}_n)). \quad (3.14)$$

Thus our Lagrangian becomes

$$\mathcal{L}_n = \sum_{n=0}^{N} G_n(\boldsymbol{x}_n, \boldsymbol{\lambda}_n, \dot{\boldsymbol{x}}_n, \dot{\boldsymbol{\lambda}}_n). \quad (3.15)$$

It is *this* discrete functional that we now want to minimise.

As for the continuous case, we require the first variation of the functional to be zero, i.e.

$$\delta \mathcal{L}_n = \delta \sum_{n=0}^{N} G_n(\boldsymbol{x}_n, \boldsymbol{\lambda}_n, \dot{\boldsymbol{x}}_n, \dot{\boldsymbol{\lambda}}_n) = \sum_{n=0}^{N} \delta G_n(\boldsymbol{x}_n, \boldsymbol{\lambda}_n, \dot{\boldsymbol{x}}_n, \dot{\boldsymbol{\lambda}}_n) = 0. \tag{3.16}$$

From (3.16) we have

$$\delta \mathcal{L}_n = \sum_{n=0}^{N} \left(\delta \boldsymbol{x}_n \frac{\partial G_n}{\partial \boldsymbol{x}_n} + \delta \dot{\boldsymbol{x}}_n \frac{\partial G_n}{\partial \dot{\boldsymbol{x}}_n} + \delta \boldsymbol{\lambda}_n \frac{\partial G_n}{\partial \boldsymbol{\lambda}_n} + \delta \dot{\boldsymbol{\lambda}}_n \frac{\partial G_n}{\partial \dot{\boldsymbol{\lambda}}_n} \right) = 0, \tag{3.17}$$

By implementing summation by parts and applying boundary conditions, leads to

$$\delta \mathcal{L}_n = \sum_{n=0}^{N} \left(\left(\frac{\partial G_n}{\partial \boldsymbol{x}_n} - \frac{d}{dt} \left(\frac{\partial G_n}{\partial \dot{\boldsymbol{x}}_n} \right) \right) \delta \boldsymbol{x}_n + \left(\frac{\partial G_n}{\partial \boldsymbol{\lambda}_n} - \frac{d}{dt} \left(\frac{\partial G_n}{\partial \dot{\boldsymbol{\lambda}}_n} \right) \right) \delta \boldsymbol{\lambda}_n \right) = 0. \tag{3.18}$$

Since $\boldsymbol{x}_n$ and $\boldsymbol{\lambda}_n$ are independent, we can thus deduce the following equations that are necessary to ensure $\delta \mathcal{L}_n = 0$,

$$\frac{\partial G_n}{\partial \boldsymbol{x}_n} - \frac{d}{dt} \left(\frac{\partial G_n}{\partial \dot{\boldsymbol{x}}_n} \right) = 0 \tag{3.19}$$

$$\frac{\partial G_n}{\partial \boldsymbol{\lambda}_n} - \frac{d}{dt} \left(\frac{\partial G_n}{\partial \dot{\boldsymbol{\lambda}}_n} \right) = 0. \tag{3.20}$$

The discrete forward model equations, as given by (3.12), are found using (3.20), whereas (3.19) provides the discrete adjoint equations. For our problem we have,

$$G_n = \sum_{n=0}^{N} (\boldsymbol{y}_n - H_n[\boldsymbol{x}_n])^T \boldsymbol{R}_n^{-1} (\boldsymbol{y}_n - H_n[\boldsymbol{x}_n]) + \boldsymbol{\lambda}_n (\dot{\boldsymbol{x}}_n - f_n(\boldsymbol{x}_n)). \tag{3.21}$$

We can thus derive the discrete adjoint equations for this problem,

$$\boldsymbol{\lambda}_N = 0 \tag{3.22}$$

$$\boldsymbol{\lambda}_n = \boldsymbol{\mathcal{F}}_n^T (\boldsymbol{\lambda}_{n+1}) - \boldsymbol{\mathcal{H}}_n^T \boldsymbol{R}_n^{-1} (\boldsymbol{H}_n(\boldsymbol{x}_n) - \boldsymbol{y}_n) \quad (n = N-1, \ldots 0), \tag{3.23}$$

where $\boldsymbol{\mathcal{F}}_n$ is the Jacobian of $f(\boldsymbol{x}_n)$, $\boldsymbol{\mathcal{H}}$ the Jacobian of $\boldsymbol{H}_n$, both with respect to $\boldsymbol{x}_n$ [40]. In practice, however, the adjoint model is derived directly from the linearisation of the forward, nonlinear model. These discrete adjoint equations are necessary to find the gradient of the observation term and, consequently, the gradient of our cost function. The minimisation problem is solved using an iterative method that requires the calculation of the cost function *and* its gradient with each iteration.

Figure 3.1 gives an illustration of the 4d Var process. We run a model trajectory from a first guess. We then measure the distance, J_b, between the background and this trajectory at the initial time only, i.e. the first guess, and the distances, J_{o_n}, between observations and the trajectory at $t = t_n$. Our cost function, J, given by (3.11), can therefore be summarised as

$$J = J_b + \sum_n J_{o_n}. \tag{3.24}$$

With each iteration we measure J_b at the initial time and the J_{o_n} across the trajectory. Our aim is to reduce these distances using a minimisation algorithm, with each iteration providing a new set of initial conditions until we arrive at the minimum. The initial conditions at the minimum give us the *analysis*, as illustrated in the second of the two diagrams in figure 3.1. The analysis therefore initialises a solution that is close to the background field and fits with the observed field.

3.3 Practical Implementation of 4d Var

We now consider the practical details of implementing the 4d Var system. In Section 3.2.2 we defined our discrete cost function in (3.11). We must minimise this in order to find our analysis.

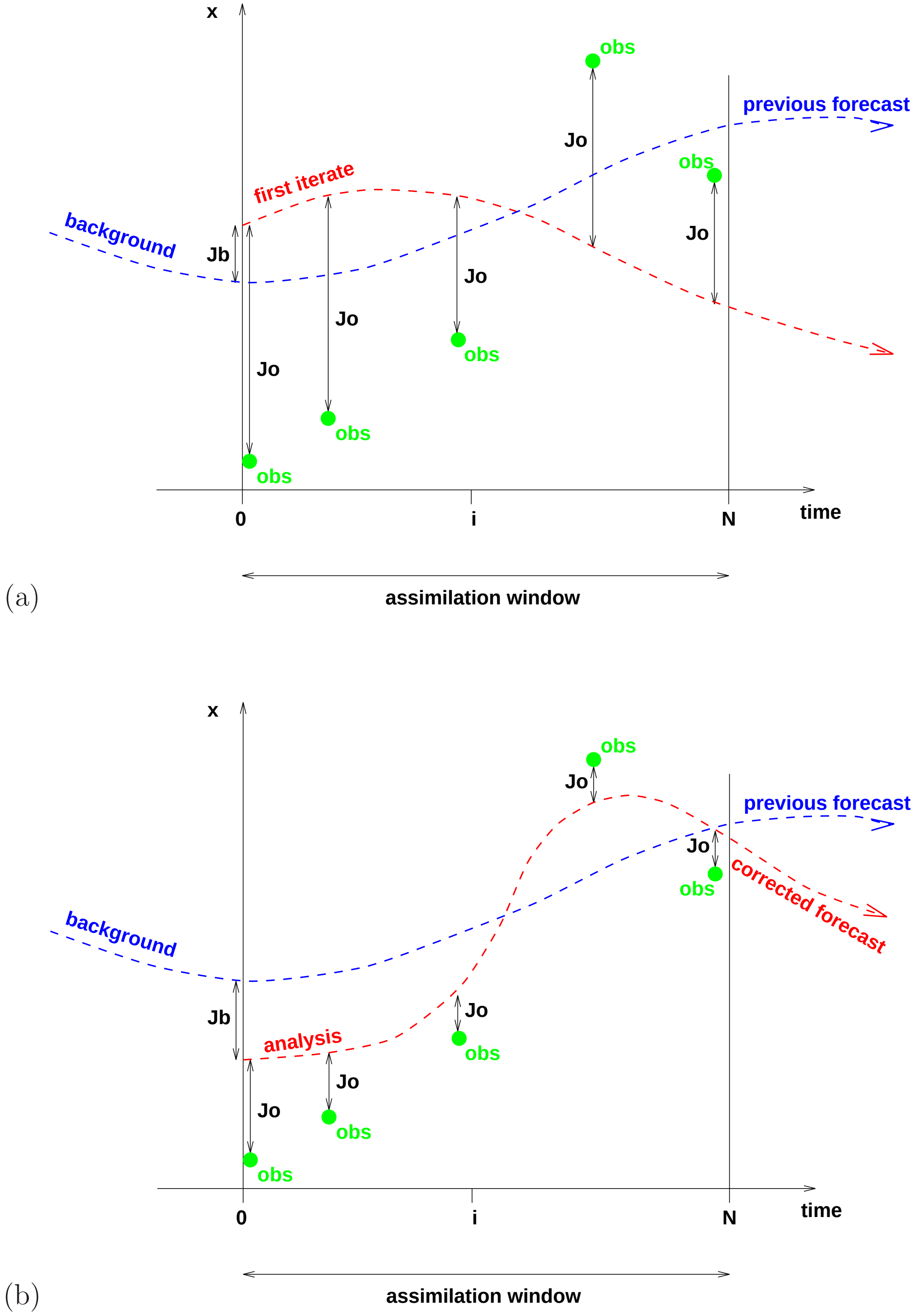


Figure 3.1: Diagrammatic representation of the 4d Var approach; (a) illustrates the trajectory given by the first guess, (b) the trajectory from the analysis.

3.3.1 Minimisation of the Cost Functional, J

The minimisation of the cost functional requires a minimisation algorithm, such as conjugate gradient or quasi-Newton methods, that can be obtained from many different sources, and so this will not be investigated here [46]. However the algorithm will require both J and its gradient ∇J and we need to consider how these can be evaluated.

The evaluation of the cost functional, J, is relatively straightforward. We need just evaluate equation (3.11) as shown. To do this we need only to run the forward non-linear model. J_b is calculated at the initial time, J_o is the sum of the J_{o_n} calculated at each time step, where

$$J_{o_n} = (\boldsymbol{y}_n - H_n[\boldsymbol{x}_n])^T \boldsymbol{d}_n \tag{3.25}$$

and $\boldsymbol{d}_n$ are the normalised departures given by $\boldsymbol{d}_n = \boldsymbol{R}^{-1}(\boldsymbol{y}_n - H_n[\boldsymbol{x}_n])$ which should be stored at each time step, as they are required in the calculation of ∇J.

We also require the gradient of the cost function given by (3.11). We find the gradient of the background term by simply finding ∇J_b. However the evaluation of $\nabla J_o = \nabla(\sum_n J_{o_n})$ is less obvious. We can write this in the following way,

$$\begin{aligned} -\frac{1}{2}\nabla J &= -\frac{1}{2}\sum_{n=0}^{N} \nabla J_{o_n} \qquad (3.26) \\ &= \sum_{n=0}^{N} \boldsymbol{M}_1^T \dots \boldsymbol{M}_n^T \boldsymbol{H}_n^T \boldsymbol{d}_n \\ &= \boldsymbol{H}_0^T \boldsymbol{d}_0 + \boldsymbol{M}_1^T[\boldsymbol{H}_1^T \boldsymbol{d}_1 + \boldsymbol{M}_2^T[\boldsymbol{H}_2^T \boldsymbol{d}_2 + \dots + \boldsymbol{M}_N^T \boldsymbol{H}_N^T \boldsymbol{d}_N] \dots], \end{aligned}$$

where $\boldsymbol{M}$ is the linearisation of the forward, nonlinear model and $\boldsymbol{H}$ is the linearised observation operator [5].

This can be evaluated from $n = N$ to $n = 0$, i.e. from right to left, that is we use an adjoint model where the adjoint is defined as the transpose of the tangent linear model. So ∇J_o is found by running the adjoint model backwards, initialising it by setting the adjoint variable $\boldsymbol{\lambda}_N = 0$ at the final time. At each backwards timestep we add the forcing term $\boldsymbol{H}_n^T \boldsymbol{d}_n$ to $\boldsymbol{\lambda}_n$, before applying the adjoint model to give $\boldsymbol{\lambda}_{n-1}$. Hence equation (3.26) becomes,

$$-\frac{1}{2}\nabla J_o = \boldsymbol{\lambda}_0 \tag{3.27}$$

3.3.2 The Tangent Linear Model

In order to derive our adjoint equations we require the tangent linear model. We must therefore develop and test the linearisation of our forward, nonlinear model. It is important that the model is tested as we must have the correct linearisation in order to find the correct adjoint. We must also show that the tangent linear hypothesis holds, otherwise the gradient will not be accurate. The tangent linear hypothesis states that the linearisation of a nonlinear model and the model itself must exhibit similar behaviour for a period of time known as the *validity time.*

To linearise our model we consider the Taylor expansion of a nonlinear model, M, around state $\boldsymbol{x}$, we thus have

$$M(\boldsymbol{x} + \delta\boldsymbol{x}) = M(\boldsymbol{x}) + M'(\boldsymbol{x})\delta\boldsymbol{x} + \frac{1}{2}M''(\boldsymbol{x})\delta\boldsymbol{x}^2 + \ldots \tag{3.28}$$

The linear model is thus given by considering only the first order terms, thus

$$M(\boldsymbol{x} + \delta\boldsymbol{x}) - M(\boldsymbol{x}) = M'(\boldsymbol{x})\delta\boldsymbol{x}, \tag{3.29}$$

and therefore our linear model is found by differentiating the nonlinear, discrete model equations with respect to the state vector. Note that we linearise the discrete equations, rather than discretise the linearised equations. The two may not necessarily produce the same set of discrete linear equations [24].

The tangent linear hypothesis says that the model can be approximated by retaining only the linear terms, neglecting terms of second order and above. Therefore if the linearisation is valid we have,

$$M(\boldsymbol{x} + \delta\boldsymbol{x}) \approx M(\boldsymbol{x}) + \boldsymbol{M}(\boldsymbol{x})\delta\boldsymbol{x}, \tag{3.30}$$

where $\boldsymbol{M}(\boldsymbol{x})$ is the linear model, defined by $\boldsymbol{M}(\boldsymbol{x}) = M'(\boldsymbol{x})$.

Thus we can see that it is vital that the assumption is shown to be valid. If the linearisation of the model does not retain the original features of the nonlinear system, then the tangent linear hypothesis is not a good assumption, and therefore the gradient will not be accurate. The final linear model can be tested in two ways.

A. Correctness Test

The first test test checks whether the linear model is the correct linearisation of the original problem in the vicinity of a given trajectory. Consider a general state variable $\boldsymbol{x}$, and perturbation $\delta\boldsymbol{x}$. From equation (3.30) we have,

$$\frac{M(\boldsymbol{x} + \delta\boldsymbol{x})_i - M(\boldsymbol{x})_i}{[\boldsymbol{M}(\boldsymbol{x})\delta\boldsymbol{x}]_i} = 1 + \text{higher order terms}, \tag{3.31}$$

where M is the nonlinear model for two different initial conditions, $\boldsymbol{M}$ is the tangent-linear model describing the evolution of the perturbation, and the subscript i denotes the i^{th} vector component. As we have seen the tangent linear model is valid only if higher order terms are negligible. Therefore if we take the limit as $\delta\boldsymbol{x}$ tends to zero of the right-hand side and subtract one then the answer should tend to zero, as shown in the following equation.

$$\lim_{\boldsymbol{x} \to 0} \frac{M(\boldsymbol{x} + \delta\boldsymbol{x})_i - M(\boldsymbol{x})_i}{[\boldsymbol{M}(\boldsymbol{x})\delta\boldsymbol{x}]_i} - 1 = 0. \tag{3.32}$$

In addition, this limit should be reached linearly, proving that the linear part of the Taylor series is explained by the tangent linear model. This is known as testing the *correctness* of the model [50]. This test is illustrated by Figure 4.7 in Chapter 4.

B. Validity Test

The second test looks at the validity of the linear model. Here we compare the trajectory of the linear model with inital conditions $\delta\boldsymbol{x}_0$, and the difference between two nonlinear model trajectories with initial condition $\boldsymbol{x}$ and $\boldsymbol{x} + \delta\boldsymbol{x}$. The validity time for the linear model is the period over which these exhibit similar behaviour. This is demonstrated by Figures 4.8 to 4.9 in Chapter 4, amongst others.

These tests are necessary to validate the linear and adjoint equations and are carried out for both of the data assimilation schemes in this thesis.

3.3.3 The Adjoint

There are two methods to construct an adjoint model.

- Constructing the discrete adjoint equations from the discrete linearised equations.
- Constructing the continuous adjoint equations then discretising.

The continous equations are those given by (3.10) and in general will not produce the same discrete model equations as the first techniques. The first of these methods can be implemented in two ways. We can use the matrix form of the discrete tangent linear equations to find the transpose or we can derive the adjoint directly from the tangent linear model code, following a set procedure for each line of code. Operationally the second of these methods is used, and that is the method we implement here. For a description of how this second method is applied see [17, 9]. Once the adjoint is constructed it can be tested in two ways.

A. Adjoint Test

We test whether the adjoint code gives the true adjoint of the tangent linear model by making use of the adjoint operator. In general we have

$$< \boldsymbol{A}\boldsymbol{x}, \boldsymbol{y} >=< \boldsymbol{x}, \boldsymbol{A}^{*}\boldsymbol{y} >, \tag{3.33}$$

where $\boldsymbol{A}$ is a linear operator, and $\boldsymbol{A}^{*}$ is its adjoint; the brackets $< ..., ... >$ denote an inner product. For our model we have the tangent linear operator $\boldsymbol{M}$, and its adjoint $\boldsymbol{M}^{T}$. We can use this equation to test our model, by first applying the forward linear model to the initial conditions to produce a final state. We then apply the adjoint to this final state. By rearranging (3.33), if the adjoint is correct then we should find,

$$< \boldsymbol{M}\delta\boldsymbol{x}_0, \boldsymbol{M}\delta\boldsymbol{x}_0 > - < \delta\boldsymbol{x}_0, \boldsymbol{M}^{T}\boldsymbol{M}\delta\boldsymbol{x}_0 >= 0. \tag{3.34}$$

B. Gradient Test

We must also test that the adjoint produces the correct gradient of the cost function. We use a Taylor expansion of the cost function [29],

$$J(\boldsymbol{x} + \delta\boldsymbol{x}) = J(\boldsymbol{x}) + \delta\boldsymbol{x}^{T}\nabla J + \dots . \tag{3.35}$$

Rearranging (3.35) we have

$$\frac{J(\boldsymbol{x} + \delta\boldsymbol{x}) - J(\boldsymbol{x})}{\delta\boldsymbol{x}^{T}\nabla J} = 1 + \dots . \tag{3.36}$$

We should therefore find that the quantity on the left hand side of (3.36) approaches 1 as $\delta\boldsymbol{x}$ approaches zero. We define $\delta\boldsymbol{x}$ by

$$\delta\boldsymbol{x} = \alpha\frac{\nabla J}{\|\nabla J\|}, \tag{3.37}$$

so that $\delta\boldsymbol{x}$ is a vector in the gradient direction and thus the variation of the variables gives a consistent scaling [29], ∇J is the gradient found using the adjoint. Using our definition of $\delta\boldsymbol{x}$ in (3.36) gives

$$\phi(\alpha) = \frac{J(\boldsymbol{x} + \alpha\frac{\nabla J}{\|\nabla J\|}) - J(\boldsymbol{x})}{\alpha\left(\frac{\nabla J}{\|\nabla J\|}\right)^T \cdot \nabla J} = 1 + O(\alpha). \tag{3.38}$$

Thus if the gradient is correct, for small values of α, ϕ should linearly approach 1 as α decreases, and the residual given by

$$\left(\frac{J(\boldsymbol{x} + \alpha\frac{\nabla J}{\|\nabla J\|}) - J(\boldsymbol{x})}{\alpha\left(\frac{\nabla J}{\|\nabla J\|}\right)^T \cdot \nabla J}\right) - 1, \tag{3.39}$$

should linearly approach zero [29].

3.3.4 The Minimisation Algorithm

There are many different minimisation algorithms that can be used for these problems. We use the Conmin routine [46]. This provides a choice of two minimisation algorithms, a conjugate gradient method and the BFGS, quasi-Newton technique. For a more complete discussion and a comparison of possible minimisation techniques see [22]. Based on work cited in [22], we will use the quasi-Newton minimisation algorithm for all our experiments.

The algorithm requires stopping criteria, these are given by

$$\frac{|J_{k-1} - J_k|}{1 + |J_k|} < \varepsilon \tag{3.40}$$

$$\frac{\|\mathbf{x}_{k-1} - \mathbf{x}_k\|}{1 + \|\mathbf{x}_k\|} < \sqrt{\varepsilon} \tag{3.41}$$

$$\frac{\|\nabla J_k\|}{1 + |J_k|} \leq \sqrt[3]{\varepsilon}. \tag{3.42}$$

All three must be satisfied in order to terminate the minimisation [18].

3.4 Weak Constraints in 4d Var

The formulation of 4d Var we have illustrated above contains a *strong* constraint. We can include further constraints on the problem by adding additional terms to the cost function. These are known as *weak* constraints and are so-called due to the fact that they do not have to be *exactly* satisfied by the solution.

3.4.1 The Continuous Problem

In (3.3) we observed that the strong constraint, given by (3.2), was multiplied by a vector of Lagrange multipliers. A weak constraint is not treated in this way. Here we add a term to the cost function in the following way [45]

$$J = \int_{t_0}^{t_1} (F(t, \boldsymbol{x}, \dot{\boldsymbol{x}}) + J_c(t, \boldsymbol{x}, \dot{\boldsymbol{x}}))\mathrm{d}t. \tag{3.43}$$

We note that J_c is an *additional* term, therefore the original strong constraint is also imposed. The weak constraint, J_c is formulated as required for the specific problem. When constructing our Euler-Lagrange equations, corresponding to (3.9) and (3.10), we must include this weak constraint. Function G becomes

$$G(\boldsymbol{x}, \boldsymbol{\lambda}, \dot{\boldsymbol{x}}, \dot{\boldsymbol{\lambda}}) = F(t, \boldsymbol{x}, \dot{\boldsymbol{x}}) + \boldsymbol{\lambda}(t)C(\boldsymbol{x}, \dot{\boldsymbol{x}}) + J_c(t, \boldsymbol{x}, \dot{\boldsymbol{x}}), \tag{3.44}$$

and thus our Euler-Lagrange equations are altered accordingly.

3.4.2 The Discrete Problem

The discrete form of the cost function given by (3.43) becomes

$$J(\boldsymbol{x}_0) = (\boldsymbol{x}_0 - \boldsymbol{x}_b)^T \boldsymbol{B}^{-1}(\boldsymbol{x}_0 - \boldsymbol{x}_b) + \sum_{n=0}^{N} \big((\boldsymbol{y}_n - H_n[\boldsymbol{x}_n])^T \boldsymbol{R}_n^{-1}(\boldsymbol{y}_n - H_n[\boldsymbol{x}_n])\big) + J_c(\boldsymbol{x}_n, \boldsymbol{x}_b). \tag{3.45}$$

The weak constraint may be imposed over the length of the data assimilation window if required. When minimising the cost function the gradient of this term must be calculated and included in the cost function gradient. The gradient must then be tested as described in Section 3.3.3.

3.4.3 Current Research

The weak constraint has been implemented for several different purposes. A method of including a term to account for the model error has been considered [39, 52]. In [39] the evolution of the forward model is given by

$$\boldsymbol{x}_{n+1} = f(\boldsymbol{x}_n) + \epsilon_n, \tag{3.46}$$

where ϵ_n denotes the model error at time t_n. The weak constraint that is added to the cost function to account for this is

$$J_c = \sum_{n=0}^{N} \epsilon_n^T Q_n^{-1} \epsilon_n, \tag{3.47}$$

where Q_n is the model error covariance matrix. The minimisation of the cost function will therefore attempt to produce an analysis with small values of the model error, ϵ_n. In [39] the addition of a weak constraint of this form was found to be effective and produced improved forecasts.

Weak constraints have also been implemented as digital filters [16, 51]. They are used to filter out the high frequency oscillations - gravity waves - that can be spuriously generated by the assimilation algorithm. In [51] this has the form

$$J_c = \beta(\boldsymbol{x}(t_{\frac{N}{2}}) - \bar{\boldsymbol{x}}(t_{\frac{N}{2}}))^T \boldsymbol{W}(\boldsymbol{x}(t_{\frac{N}{2}}) - \bar{\boldsymbol{x}}(t_{\frac{N}{2}})), \tag{3.48}$$

where β is a parameter that determines how strongly the constraint will be enforced, $\boldsymbol{W}$ is a diagonal matrix that determines the weighting of each variable and

$$\bar{\boldsymbol{x}}(t_{\frac{N}{2}}) = \sum_{n=0}^{N} \alpha_n \boldsymbol{x}(t_n). \tag{3.49}$$

The latter variable, $\bar{\boldsymbol{x}}$, is the filtered model state which is defined at the middle of the time window and α_n is a weighting function. Thus the constraint term measures the distance between the unfiltered and filtered states and the minimisation of the cost function seeks to reduce this. The analysis is therefore constrained to be close to this filtered state. In both [16] and [51] the addition of such a term is found to be effective in reducing these spurious high frequency waves and produces a better forecast.

A further example of a weak constraint also aims to reduce the spurious high frequency waves but not by the use of a digital filter. In [14], Dixon and Roulstone consider using such a constraint to include more information about the system. Here they incorporate a balance condition that could potentially allow some realistic high frequency oscillations to exist in the solution. They consider the balanced equation system that is constructed from the shallow water equations [35]. The constraint has the form

$$J_c = w_{J_c} \frac{1}{2} \frac{1}{\sum_i a_i} \sum_i a_i \left(\frac{\partial \delta'(0)}{\partial t} \right)^2, \tag{3.50}$$

where w_{J_c} is the weighting of the constraint, the summation is over all the grid points in the model space and $\delta'(0)$ is the incremental divergence field at the initial time. They define

$$a_i = (\cos(\phi) * \Delta p)|_i, \tag{3.51}$$

to be the mass and area weighting of the ith grid point, where p is the pressure and ϕ is the latitude. This constraint aims to reduce the spurious gravity waves by imposing incremental non linear balance on the system. This constraint was found to be unsuccessful in that the incremental nonlinear balance was not imposed correctly. However further formulations discussed in the paper proved more successful and further work in this direction is in progress.

The last of the three weak constraints discussed here is the closest to the formulation that we will consider in this thesis. We aim to make use of the inherent characteristics of the dynamical system when imposing a further constraint.

3.5 Summary

In Chapter 3 we have introduced the idea of data assimilation. We have discussed the important features and basic concepts of a successful data assimilation scheme. Starting with the variational approach we have detailed the mathematical background of 4d Var. The linear algebra formulation has also been described since it is this approach that is used in the practical implementation of this assimilation technique.

In Section 3.3 we considered the construction of a 4d Var algorithm. We have seen that it is necessary to derive the linear model and corresponding adjoint for the problem. These must both be tested. The linear model should be a good approximation to the nonlinear model for the length of the assimilation window and the adjoint must give the correct gradient of the cost function at the optimal solution.

In the final section we discussed the addition of weak constraints to the cost function. We briefly illustrated the form these take before discussing current research that has made use of this formulation of 4d Var.

In the next chapter we make use of the theory discussed both here and in

Chapter 2. We consider the construction of a 4d Var algorithm for a simple Hamiltonian system. We use this to carry out a series of experiments and investigate the effect of adding a weak constraint to the scheme.

Chapter 4

System I: The Two-Body Problem

In Chapter 2 we derived the canonical equations for Hamiltonian systems and considered the special mathematical characteristics of such problems. In particular we highlighted the conservation properties of these systems and discussed how in many cases the conserved quantity, the Hamiltonian, is the total energy of the system. We also described the numerical methods that are best suited to these systems, with specific reference to symplectic schemes.

In Chapter 3 we introduced the concept of data assimilation and discussed in detail the method of 4d Var. This involves the minimisation of a cost function measuring the distance between the solution and a first guess and the solution and the observations. Also in Chapter 3 we considered the addition of weak constraints to the 4d Var cost function and looked at how and why these might be used.

In this chapter we use the information from these introductory chapters in order to construct a 4D Var algorithm for the two-body problem. We begin by deriving the canonical equations for the two-body problem. We then consider the construction of a nonlinear, forward model for this simple Hamiltonian system and investigate both its accuracy and conservation properties. We follow this by detailing the stages required to create a 4d Var algorithm using this numerical model, including the construction and testing of the corresponding linear and adjoint models. Finally we carry out a series of identical-twin experiments using a variety of observations and including additional weak constraints making use of the characteristics inherent within a Hamiltonian system. We want to demonstrate that the addition of these extra constraints will improve the analysis and provide a better forecast.

4.1 The Two-Body Problem

In this section we discuss the two-body problem, a simple example of a Hamiltonian system. We consider its conservation properties and derive the Hamiltonian form of the equations of motion.

The two-body problem describes the motion of two bodies, with masses m_1 and m_2, in mutual orbit. In its most general form this gives the trajectories of the two bodies as they move through space. This can be simplified by introducing the idea of the *reduced* two-body problem. By restating the problem in terms of the position of the centre of mass of the two bodies and the separation between them, we can manipulate Newton's equations of motion such that the full two-body problem becomes an equivalent one body problem [23]. If we consider the system in a *centre of mass frame* we have a body of reduced mass $\mu = \frac{m_1 m_2}{m_1 + m_2}$, orbiting a fixed body of mass $M = m_1 + m_2$, and thus we have motion restricted to a plane.

This system is much simpler to solve, and in addition obeys a series of rules known as Kepler's Laws. In the early seventeenth century Kepler developed these to describe the motion of the planets in the solar system. This is clearly not a reduced system. However if we consider two bodies, such as the sun and Earth, with masses m_1 and m_2 where $m_1 \gg m_2$, the reduced system has a reduced mass $\mu \approx m_2$ orbiting total mass $M \approx m_1$. Thus, in this case, the reduced problem and the full system are very similar and therefore Kepler's qualitative laws apply to the reduced problem.

Kepler's laws give a useful insight into the behaviour of the reduced two-body system and will help our understanding of later results. These laws are

- The planets follow an elliptical orbit, with the sun at one focus.
- A line from the planet to the sun will sweep out equal areas in equal times.
- The square of the period of the orbit is proportional to the cube of its semi-major axis.

These laws arose from observation and trial and error. They are, however, simply a result of Newton's three laws of motion and his law of universal gravitation. They can be restated for any two-body system as [42]

- The shape of an orbit is a conic section with the centre of mass at one focus.
- Orbital motions conserve angular momentum.
- $T_o^2 = \frac{4\pi^2}{G(M_1+M_2)}a^3$, where T_o is the orbital period, a the semi-major axis, G the universal gravitational constant and M_1, M_2 the masses of the two bodies.

We can thus see that, in addition to being a Hamiltonian system, the two-body problem has further important qualitative characteristics such as conservation of angular momentum.

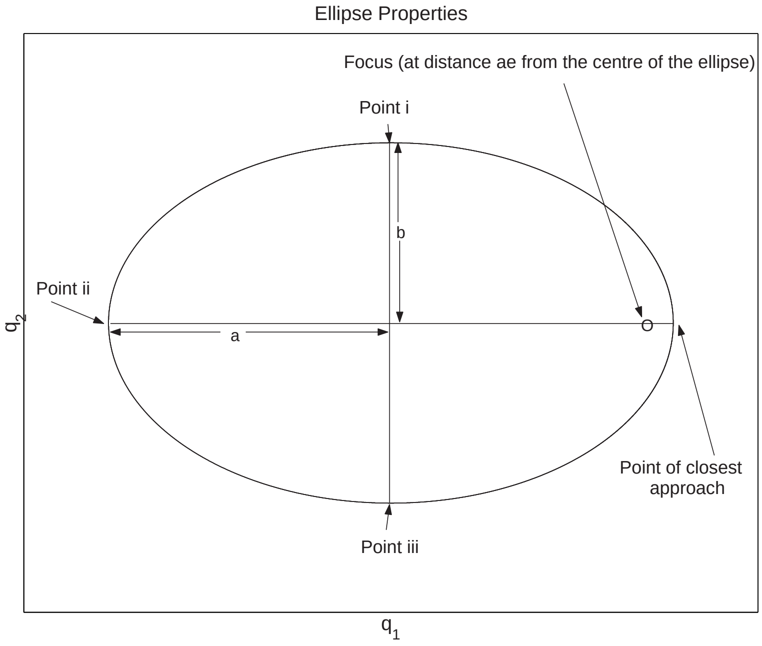


Figure 4.1: Properties of an ellipse

In this chapter we consider a reduced mass problem with a closed orbit and thus the conic section described by the first law is an ellipse. This elliptical orbit can be described using a quantity known as the eccentricity, e. If $e = 0$ then the resulting trajectory is a circular orbit. For $0 < e < 1$ we have an ellipse, and e gives a measure of the elongation of the resulting ellipse. The eccentricity is defined as

$$e = \left(1 - \frac{b^2}{a^2}\right)^{\frac{1}{2}}, \tag{4.1}$$

where a and b are the length of the semi-major and semi-minor axes of the ellipse respectively, as illustrated in Figure 4.1. The focus lies at a distance ae from the centre of the ellipse.

The reduced problem we are considering can be further simplified by introducing non-dimensional variables. Say we have initial conditions, $\boldsymbol{q}_0$, for the postion of the reduced body given by

$$\boldsymbol{q}_0 = (d, 0), \tag{4.2}$$

where d is the initial distance between the reduced mass and the centre of mass. We weight the position variables by distance, d, to give our new variables

$$\tilde{\boldsymbol{q}} = \frac{\boldsymbol{q}}{d}, \tag{4.3}$$

which is equivalent to measuring distance in units of d. In addition time is measured in units of $\left(\frac{d^3}{GM}\right)^{\frac{1}{2}}$ and the transformation to the non-dimensional time is thus given by

$$\tilde{t} = \left(\frac{GM}{d^3}\right)^{\frac{1}{2}} t, \tag{4.4}$$

where $M = m_1 + m_2$ is the total mass and G is the universal gravitational constant. The non-dimensional masses are given by

$$\tilde{m}_i = \frac{m_i}{M}. \tag{4.5}$$

In addition we have the non-dimensional momentum, $\tilde{\boldsymbol{p}}$, measured in units of $\left(\frac{G}{Md}\right)^{\frac{1}{2}}$. For clarity, as we will only be using this non-dimensional system in the thesis, we will drop the $\sim$ from these variables.

4.1.1 The Continuous Problem

As we have discussed the system is confined to a plane. We therefore describe the system using coordinates $\boldsymbol{q} = (q_1, q_2)$ and $\boldsymbol{p} = (p_1, p_2)$, which give the position and momentum, respectively, of the orbiting body. To derive the Hamiltonian form of the continuous equations of motion we need to know the Hamiltonian. As the two-body problem is a conservative system we know that this is given by the sum of the kinetic and potential energy of the system [2, 26]. The kinetic energy is given by

$$T(\boldsymbol{p}) = \frac{1}{2}\left(p_1^2 + p_2^2\right), \tag{4.6}$$

while the potential energy is

$$V(\boldsymbol{q}) = -\frac{1}{(q_1^2 + q_2^2)^{\frac{1}{2}}}, \tag{4.7}$$

since we are working in the non-dimensional system. We know from (2.18) that if the Lagrangian does not depend explicitly on t then the Hamiltonian is conserved. Since for this problem

$$\mathcal{L} = T - V = \frac{1}{2}\left(p_1^2 + p_2^2\right) + \frac{1}{(q_1^2 + q_2^2)^{\frac{1}{2}}} \tag{4.8}$$

we can see that this is true and thus the Hamiltonian is conserved for this problem.

The Hamiltonian is therefore

$$H(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2}\left(p_1^2 + p_2^2\right) - \frac{1}{(q_1^2 + q_2^2)^{\frac{1}{2}}}. \tag{4.9}$$

Using equations (2.25) and (2.26), we can thus deduce the canonical form for this system is given by

$$\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} = \boldsymbol{p} \tag{4.10}$$

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -\frac{\boldsymbol{q}}{\|\boldsymbol{q}\|^3}, \tag{4.11}$$

where $\|\boldsymbol{q}\| = (\boldsymbol{q} \cdot \boldsymbol{q})^{\frac{1}{2}}$.

These equations describe the motion of the reduced mass as it follows an elliptical orbit around the fixed centre of mass. It is these equations that we shall later discretise to provide our numerical model.

4.1.2 Conservation Properties

As we have discussed in the previous section this two-body system conserves the Hamiltonian. For this reduced, non-dimensional problem this is the total energy, E, of the system, given by

$$H(\boldsymbol{q}, \boldsymbol{p}) = E(\boldsymbol{q}, \boldsymbol{p}) = \frac{1}{2}\left(p_1^2 + p_2^2\right) - \frac{1}{(q_1^2 + q_2^2)^{\frac{1}{2}}}. \tag{4.12}$$

In addition we recall that the second of Kepler's laws stated that a line joining the orbiting body with a focus of the ellipse will sweep out equal areas in equal times. This is a result of the conservation of angular momentum by the system and thus we have a second conservation law. For our system this is given by

$$L = q_1 p_2 - p_1 q_2. \tag{4.13}$$

4.1.3 The Discrete Problem

As we have discussed in section 2.2.1, geometric methods are the most suitable approach for modelling Hamiltonian problems. We will implement the Störmer-Verlet method discussed in section 2.2.2. This symplectic scheme is given by equations (2.39) to (2.41).

If we consider the Hamiltonian for this problem given by (4.9), we can see that is separable, i.e. has the form

$$H(\boldsymbol{p}, \boldsymbol{q}) = H_1(\boldsymbol{p}) + H_2(\boldsymbol{q}). \tag{4.14}$$

where,

$$\begin{aligned} H_1(\boldsymbol{p}) &= \frac{1}{2}\left(p_1^2 + p_2^2\right) && (4.15) \\ H_2(\boldsymbol{q}) &= -\frac{1}{(q_1^2 + q_2^2)^{\frac{1}{2}}}. && (4.16) \end{aligned}$$

Using equations (2.39) to (2.41) we can derive our discrete model equations for the two-body system,

$$\begin{aligned}
\boldsymbol{P}^{n+\frac{1}{2}} &= \boldsymbol{P}^{n} - \frac{h}{2}\frac{\boldsymbol{Q}^{n}}{(Q_1^{n^2} + Q_2^{n^2})^{\frac{3}{2}}} && (4.17)\\
\boldsymbol{Q}^{n+1} &= \boldsymbol{Q}^{n} + h\boldsymbol{P}^{n+\frac{1}{2}} && (4.18)\\
\boldsymbol{P}^{n+1} &= \boldsymbol{P}^{n+\frac{1}{2}} - \frac{h}{2}\frac{\boldsymbol{Q}^{n+1}}{(Q_1^{n+1^2} + Q_2^{n+1^2})^{\frac{3}{2}}}, && (4.19)
\end{aligned}$$

where the discrete variables are given by the upper case letters. As we have discussed in section 2.2.2, each of the three steps given by (4.17) to (4.19) forms a distinct Hamiltonian system. Thus each one forms a symplectic mapping from one time step to the next, and the whole system is therefore symplectic. As we know from section 2.2.1, this method, instead of approximating the solution of the continuous system, will exactly solve an approximate Hamiltonian system.

4.2 Modelling the Two-Body Problem

In this section we will discuss the implementation of the discrete model and demonstrate that the method we have chosen is an appropriate choice for this problem. In section 4.1 we showed that the problem of two bodies in mutual orbit can be reduced to that of one body of reduced mass, μ, orbiting a fixed centre with total mass $M = m_1 + m_2$, and that in addition we can further simplify the problem by introducing non-dimensional variables. We begin here by considering the *trajectory* of the orbiting body given by the numerical model, defined solely by the initial conditions. We use the following initial conditions [8] for position, $\boldsymbol{Q}$, and momentum, $\boldsymbol{P}$, which correspond to the body starting from the perihelion, the point of closest approach. The resulting orbit then has a period of 2π, with semi-major axis, $a = 1$.

The initial conditions are

$$Q_1 = 1 - e \tag{4.20}$$

$$Q_2 = 0 \tag{4.21}$$

$$P_1 = 0 \tag{4.22}$$

$$P_2 = \sqrt{\frac{1+e}{1-e}}, \tag{4.23}$$

where e is the eccentricity of the resulting ellipse. Note that the period, T_o, and semi-major axis, a are independent of the value of e we select. In addition the initial conditions are such that our choice of e is sufficient to describe the trajectory for all time. For all our experiments with the two-body problem, we run the model with a timestep of $h = 0.001$ unless stated otherwise.

The analytic solution to this problem would result in a trajectory that follows the same path on each orbit. We therefore consider the trajectories produced by our numerical model for three different values of the eccentricity, $e = 0$, $e = 0.5$, and $e = 0.9$. In each case we run the model for $t \approx 20\pi$ corresponding to approximately ten orbits. Figure 4.2 shows the evolution of the orbiting body for each value of e. We see that in all three cases the trajectory does not deviate from the initial orbital path. This suggests that our numerical model, using the Störmer-Verlet symplectic method, is a good choice for this problem.

We can further test the model by looking at the behaviour of the *total energy* of the system. As we have discussed the two-body problem is Hamiltonian and therefore has inherent conservation properties. From section 4.1.2 we know that the Hamiltonian for this system is equivalent to the total energy and ideally this should be conserved by our choice of numerical scheme. In section 2.2.1 we stated that *symplectic* schemes preserved the conservation properties of a system. We

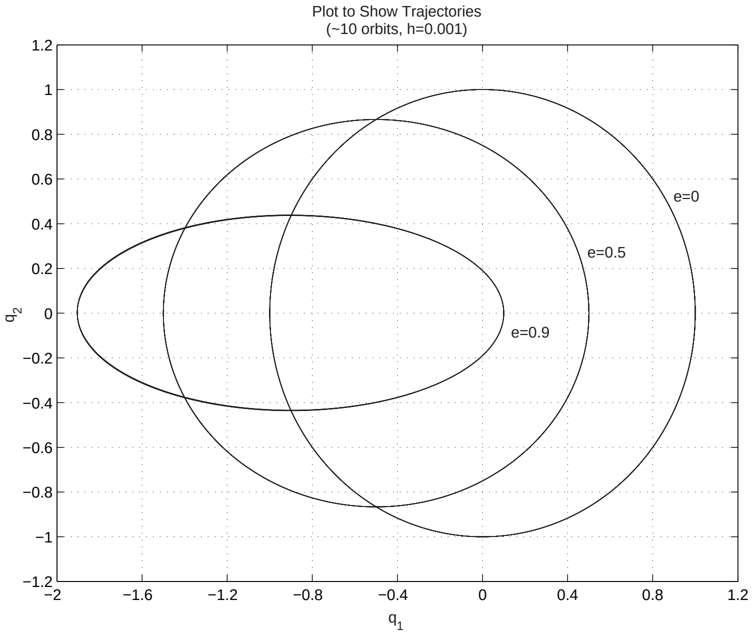


Figure 4.2: Trajectories given by the Störmer-Verlet method for $e = 0$, $e = 0.5$, $e = 0.9$. Timestep, $h = 0.001$.

therefore expect our scheme will conserve the total energy of the system and will thus be a good choice for our investigations into the two-body problem.

To illustrate this we first calculate the true value of the energy for our set of initial conditions. From (4.12) we know the total energy, equivalent to the Hamiltonian, is

$$E = \frac{1}{2}\left(p_1^2 + p_2^2\right) - \frac{1}{(q_1^2 + q_2^2)^{\frac{1}{2}}}. \tag{4.24}$$

Substituting the initial conditions given by (4.20) to (4.23) into (4.24), the energy of this system is

$$E = \frac{1}{2}\left(\sqrt{\frac{1+e}{1-e}}\right)^2 - \frac{1}{\sqrt{(1-e)^2}} = -\frac{1}{2} \tag{4.25}$$

and is therefore independent of our choice of eccentricity. If energy is conserved by the numerical scheme then our model should produce a trajectory for the orbiting body with energy, $E = -0.5$ for the full length of the model run. Thus the difference between this value and that given by the model should be zero.

Again we run our model for three different values of the eccentricity, $e = 0$, $e = 0.5$, and $e = 0.9$. Figure 4.3 shows, for three values of e, the difference between the true energy, calculated analytically from the initial conditions, and the energy given by the model at each timestep. From the plots, we can see that the error in the numerical solution for the case $e = 0$, while not constant, is at least only a very small deviation, of order 10^{-13}, from the zero value that the exact solution would give. As the eccentricity increases, so does this deviation. For $e = 0.5$ we can see that the vertical scale is 10^{-6} and for $e = 0.9$ the deviation is of order 10^{-3}.

We can also observe that this energy difference for the last two cases includes some sharp troughs. These begin at $t = 0$ and repeat at intervals of 2π. This corresponds with the body passing the point of closest approach - the perihelion. Note there are no sharp deviations in the case of a circular orbit where $e = 0$ as all points on the orbit are equidistant from the focus. We can explain these sharp troughs in the model error by considering the second of Kepler's Laws dicussed in section 4.1. This states that a line between the focus and the orbiting body will sweep out equal areas in equal times. Thus when the body is close to the focus it must travel faster than when it is further away. As we are using a fixed step scheme for our numerical model this implies that the area near the perihelion is modelled by fewer steps, and we would thus expect it to be less accurate in this region.

This is confirmed if we observe what happens to the model error in the energy in this region if we reduce the stepsize, as shown in figure 4.4. Again the eccentricity is 0.9 and here we compare the stepsizes, h=0.001 and h=0.0005.

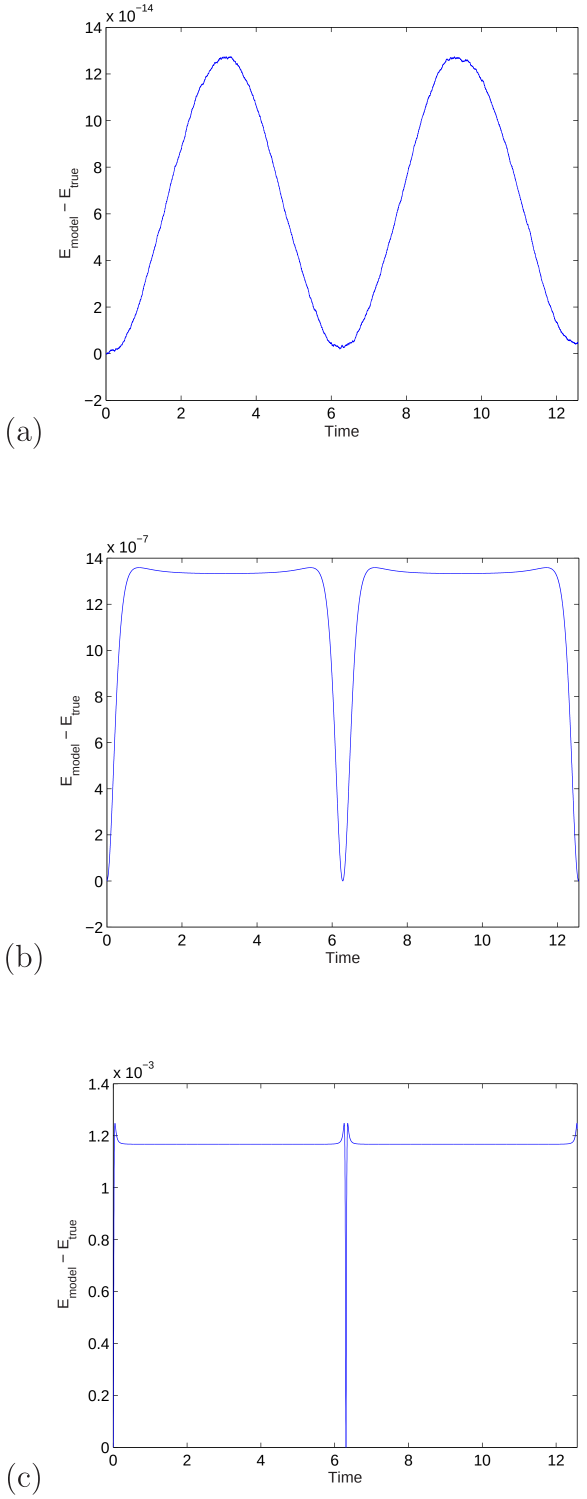


Figure 4.3: Difference between the true energy (calculated analytically) and that given by the Störmer-Verlet model for (a) $e = 0$, (b) $e = 0.5$ and (c) $e = 0.9$, timestep, h=0.001

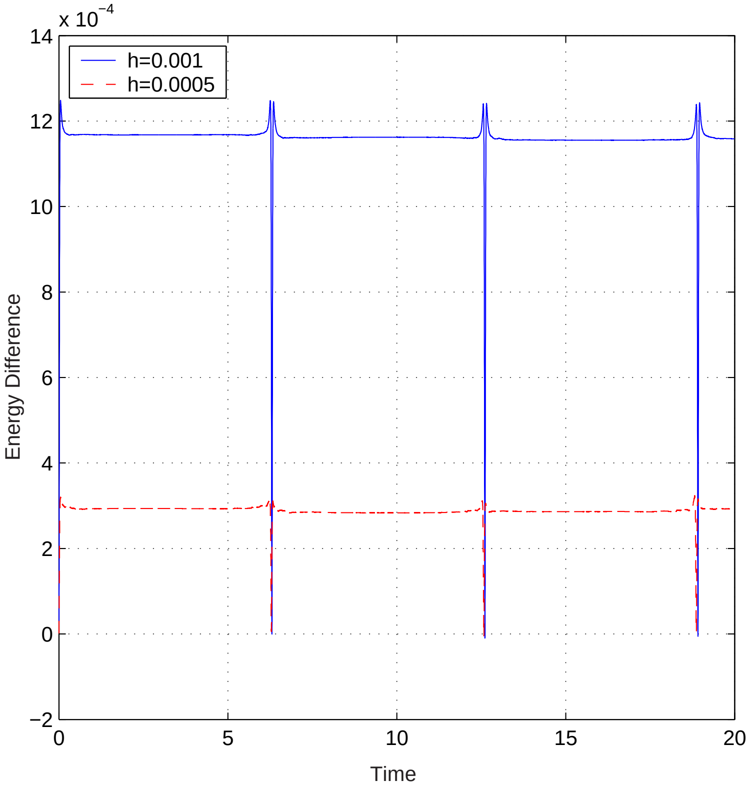


Figure 4.4: Difference between the true energy and the model energy for two different values of h, $h = 0.0001$ and $h = 0.001$, eccentricity $e = 0.9$

We can see that the difference between the model energy and the true value is considerably smaller for the reduced stepsize, and the size of the troughs is reduced at the perihelion. Figure 4.4 can also be used to demonstrate that our model has been coded correctly. If we consider the error in the energy in this figure, we can see that when the step size is halved, this error is reduced to approximately a quarter of the original value. This is as expected for a second order model such as the method we use here.

We again consider Figure 4.3(c). Here we observe that the energy given by the model in the sharp troughs is actually closer to the true value than at other points in the orbit. This is because we are using initial conditions that correspond to the perihelion. In this case we are therefore imposing the true value of the energy at the point of closest approach. We can change the initial conditions so that they correspond to another point on the orbit. For example for the aphelion, the furthest point, the initial conditions are

$$\boldsymbol{Q}(t_0) = (-(1+e), 0) \tag{4.26}$$

$$\boldsymbol{P}(t_0) = \left(0, -\sqrt{\frac{1-e}{1+e}}\right) \tag{4.27}$$

Figure 4.5 illustrates the difference in energy between the model and the truth when we use these new initial conditions. Here we see that again the sharp troughs correspond to the points of closest approach, but this time the energy difference is greater at this point.

Our results show that where we have low values of the eccentricity our model produces excellent results both in terms of the trajectory and the conservation of energy. However we have seen that where the eccentricity is higher the model, while still producing a good trajectory, does not do quite so well with modelling the energy due to the fixed timestep.

To compensate for this inaccuracy at the perihelion it is possible to apply a variable step method. Often this adaptivity can be produced by placing a bound on some specified estimation of local accuracy [47]. However this is not appropriate when using geometric integration methods, as it does not consider the underlying structure of the system. The adaptive Verlet method [21, 28] uses a different approach. Here we introduce a time-regularisation, such that the evolution evolves in *fictive time* τ [7]. In this way regions which have a faster variation with time, such as we see at the perihelion, vary more slowly with τ and thus the errors we previously observed should be reduced. This fictive time is introduced via a

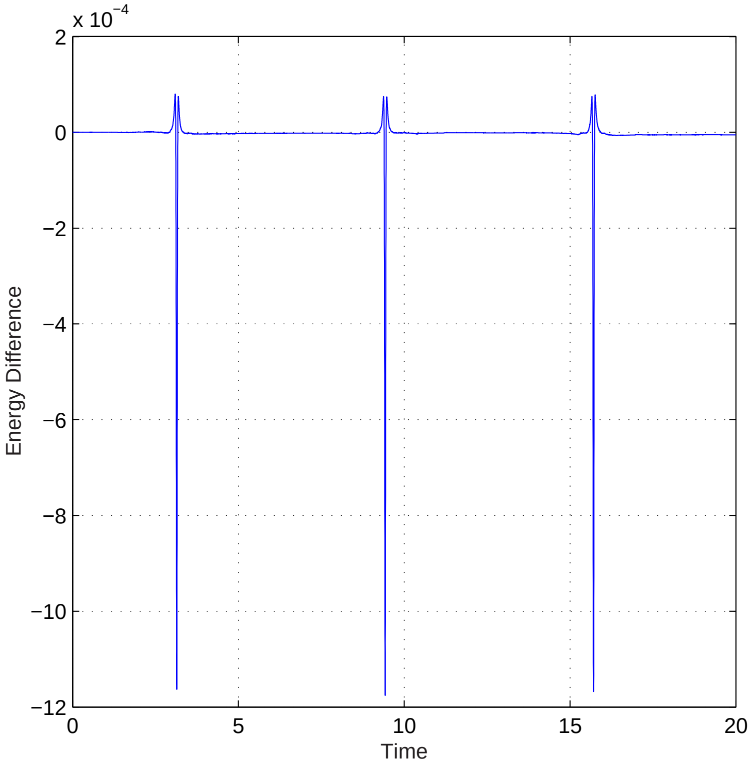


Figure 4.5: Difference between the true energy and the model energy for a trajectory starting at the aphelion, eccentricity $e = 0.9$, stepsize h=0.001

Sundman transformation,

$$\frac{dt}{d\tau} = g(\boldsymbol{q}, \boldsymbol{p}), \tag{4.28}$$

where g is a scalar function, such that g is small when the system is more rapidly evolving, and is chosen to suit the problem in question. Thus our system becomes,

$$\frac{d\boldsymbol{p}}{d\tau} = -g\left(\frac{\boldsymbol{q}}{(q_1^2 + q_2^2)^{\frac{3}{2}}}\right) \tag{4.29}$$

$$\frac{d\boldsymbol{q}}{d\tau} = g\boldsymbol{p}. \tag{4.30}$$

For this two-body problem a suitable choice for g, given by [6], is

$$g = \boldsymbol{r}^{\frac{3}{2}}, \tag{4.31}$$

where $\boldsymbol{r} = (q_1^2 + q_2^2)^{\frac{1}{2}}$.

Our discrete equations are

$$
\begin{aligned}
\boldsymbol{Q}^{n+\frac{1}{2}} &= \boldsymbol{Q}^n + \frac{\Delta\tau}{2\rho^n}\boldsymbol{P}^n && (4.32)\\
\rho^{n+1} &= \frac{2}{g(\boldsymbol{Q}^{n+\frac{1}{2}})} - \rho^n && (4.33)\\
\boldsymbol{P}^{n+1} &= \boldsymbol{P}^n - \frac{\Delta\tau}{2}\{\frac{1}{\rho^n} + \frac{1}{\rho^{n+1}}\}\frac{\boldsymbol{Q}^{n+\frac{1}{2}}}{({Q_1^{n+\frac{1}{2}}}^2 + {Q_2^{n+\frac{1}{2}}}^2)^{\frac{3}{2}}} && (4.34)\\
\boldsymbol{Q}^{n+1} &= \boldsymbol{Q}^{n+\frac{1}{2}} + \frac{\Delta\tau}{2\rho^{n+1}}\boldsymbol{P}^{n+1} && (4.35)\\
t^{n+1} &= t^n + \frac{\Delta\tau}{2}\{\frac{1}{\rho^n} + \frac{1}{\rho^{n+1}}\}. && (4.36)
\end{aligned}
$$

The initial conditions are

$$
\begin{aligned}
\boldsymbol{Q} &= (-e, \sqrt{1-e^2}) && (4.37)\\
\boldsymbol{P} &= (-1, 0) && (4.38)\\
\rho &= \frac{1}{g(t_0)} && (4.39)
\end{aligned}
$$

We now consider whether this method models the energy more effectively at high eccentricities. Figure 4.6 shows the results for $e = 0.9$. We see that there has been an improvement to the solution. The difference between the model energy and the true energy is now of order 10^{-6}. This is three orders of magnitude smaller than the results given using the Störmer-Verlet scheme.

However the changes made to introduce this adaptivity mean that the system modelled is no longer Hamiltonian. While it is possible to make changes to the scheme to return to a Hamiltonian regime [3], this results in a complex numerical method. As we are looking for a simple model for our data assimilation scheme we will use the Störmer-Verlet method for all further work. We believe that the problems encountered with the energy at the point of closest approach will not

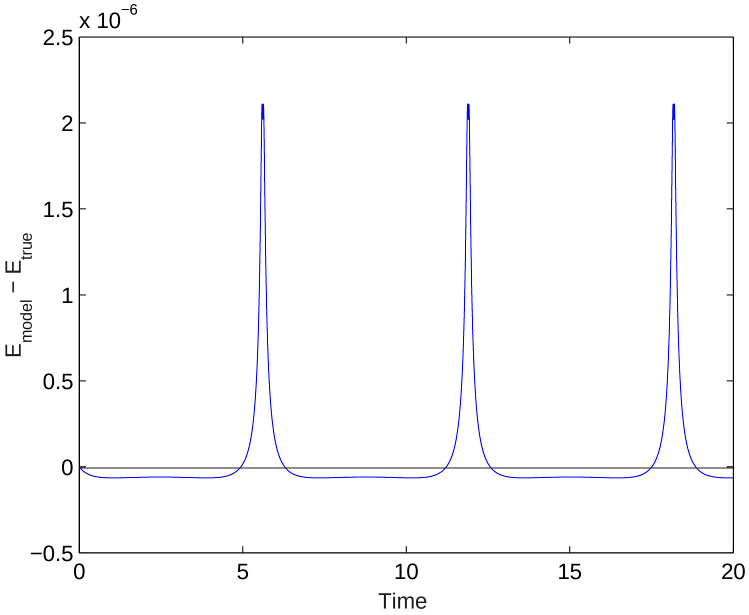


Figure 4.6: Difference between the true energy and the model energy given by the adaptive Verlet method, eccentricity $e = 0.9$, variable stepsize

prove to be significant. The energy difference always returns to the same value after the trough due to the symmetry of the problem, thus there is no loss or gain of energy from the system.

4.3 Developing 4d Var for the Two-Body System

As we explained in Chapter 3, in order to construct our data assimilation scheme for the two-body problem we need several ingredients; a forward nonlinear model and its linearisation in order to derive the adjoint, along with a suitable minimisation algorithm. In the previous section we discussed the first of these - a nonlinear model for the two-body problem - we will now consider the remaining three.

4.3.1 The Linear Model

To produce the linear model of the two-body problem we must linearise the discrete nonlinear equations given by (4.17) to (4.19). Following the procedure outlined in Chapter 3 our linear equations are given by (4.40) to (4.42). Here $\delta\boldsymbol{Q}^n$ and $\delta\boldsymbol{P}^n$ are the position and momentum perturbations respectively at time step n. The vectors $\boldsymbol{Q}^n$ and $\boldsymbol{P}^n$ are the linearisation states, found using the unperturbed nonlinear model.

$$\begin{aligned}
\delta\boldsymbol{P}^{n+\frac{1}{2}} &= \delta\boldsymbol{P}^n - \frac{h}{2}\frac{\delta\boldsymbol{Q}^n}{(Q_1^{n^2}+Q_2^{n^2})^{\frac{3}{2}}} + \frac{h}{2}\frac{3\boldsymbol{Q}^n(Q_1^n\delta Q_1^n + Q_2^n\delta Q_2^n)}{(Q_1^{n^2}+Q_2^{n^2})^{\frac{5}{2}}} && (4.40)\\
\delta\boldsymbol{Q}^{n+1} &= \delta\boldsymbol{Q}^n + h\delta\boldsymbol{P}^{n+\frac{1}{2}} && (4.41)\\
\delta\boldsymbol{P}^{n+1} &= \delta\boldsymbol{P}^{n+\frac{1}{2}} - \frac{h}{2}\frac{\delta\boldsymbol{Q}^{n+1}}{(Q_1^{n+1^2}+Q_2^{n+1^2})^{\frac{3}{2}}}\\
&\quad + \frac{3\boldsymbol{Q}^{n+1}(Q_1^{n+1}\delta Q_1^{n+1} + Q_2^{n+1}\delta Q_2^{n+1})}{(Q_1^{n+1^2}+Q_2^{n+1^2})^{\frac{5}{2}}}. && (4.42)
\end{aligned}$$

The linearisation needs to be tested in two different ways as described previously in Chapter 3.

A. Correctness Test

In order to test this linear *code*, we need to check that the tangent-linear model is the correct linearisation of the non-linear model in the vicinity of a given trajectory. From (3.32) we know that we need to calculate

$$\psi(\delta\boldsymbol{x}) = \frac{M(\boldsymbol{x}+\delta\boldsymbol{x}) - M(\boldsymbol{x})}{L\delta\boldsymbol{x}} - 1, \tag{4.43}$$

where $\boldsymbol{x} = (\boldsymbol{Q}, \boldsymbol{P})$, $\delta\boldsymbol{x} = (\gamma\boldsymbol{Q}, \gamma\boldsymbol{P})$, M is the nonlinear model and L is the linear model. We find ψ for a range of γ.

We then plot $||\psi(\delta\boldsymbol{x})||$ using the $L2$ norm, versus γ, as shown in Figure 4.7. From this test we can see that as $\gamma \to 0$ and therefore $\delta\boldsymbol{x} \to 0$, $||\psi(\delta\boldsymbol{x})|| \to 0$, and

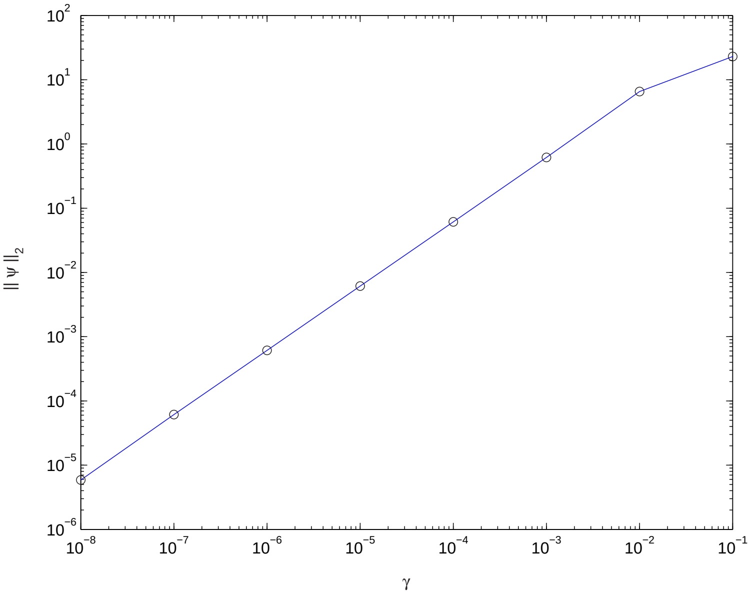


Figure 4.7: Graph to illustrate the correctness of the two-body linear model, $e = 0$, final time$= 2\pi$.

this limit is reached linearly. As discussed in Section 3.30 we can thus conclude that the tangent linear model code is correct.

B. Validity Test

The test for correctness implies that the linear code is correct; however, this does not necessarily mean that the tangent linear hypothesis holds for this problem. To test this we must consider the *validity* of the linear model. As we have discussed in Section 3.30B the validity is the length of time for which the linear model is a good approximation to the non-linear problem. There are of course errors associated with linearisation, and these will increase as time progresses. The model is valid only where the linear model mirrors the behaviour of the original problem to a reasonable degree of accuracy. To do this, we need to track the evolution of a perturbation in both of our models.

We run the non-linear model with two sets of initial conditions, $(\boldsymbol{Q}_0, \boldsymbol{P}_0)$ and $((1+\gamma)\boldsymbol{Q}_0, (1+\gamma)\boldsymbol{P}_0)$, where vectors $\boldsymbol{Q}_0$ and $\boldsymbol{P}_0$ are the same initial conditions as listed in equations (4.20) and (4.23), and the value of γ can be varied as required. The tangent linear model uses the data from the unperturbed run of the non-linear model as the linearisation state, and is initialised using initial conditions $(\gamma\boldsymbol{Q}_0, \gamma\boldsymbol{P}_0)$. We then compare the result of the tangent linear model, which returns values for $\delta\boldsymbol{Q}$ and $\delta\boldsymbol{P}$ at each time step, with the difference in position and momenta at each step produced by the two model runs of the non-linear scheme. The validity time is thus the point at which the non-linear and linear model results begin to separate greatly.

This test has been done for various sizes of γ and eccentricity. Figures 4.8 and 4.9 show for each case the L2 norm of the perturbation, $||\delta\boldsymbol{x}||_2$, where $\delta\boldsymbol{x} = (\delta\boldsymbol{Q}, \delta\boldsymbol{P})$.

Figure 4.8 illustrates the effect of changing the size of the perturbation, γ, whilst keeping the eccentricity fixed at $e = 0$. We show the results for $\gamma = 10^{-3}$, $\gamma = 10^{-2}$ and $\gamma = 10^{-1}$ respectively. These graphs confirm that the validity time is longer for smaller perturbations. This is as we would expect since the smaller the size of the perturbation, the smaller the size of the non-linear components that have been neglected in the linear model.

Figure 4.9 shows how changing the eccentricity affects the validity, whilst fixing $\gamma = 10^{-3}$. From these we see that the validity time is shorter for larger eccentricities. For $e = 0$, the two models exhibit similar behaviour, and the validity time is long. However for $e = 0.5$, the behaviour of the linear model gradually becomes less consistent with the behaviour of the nonlinear model. This difference

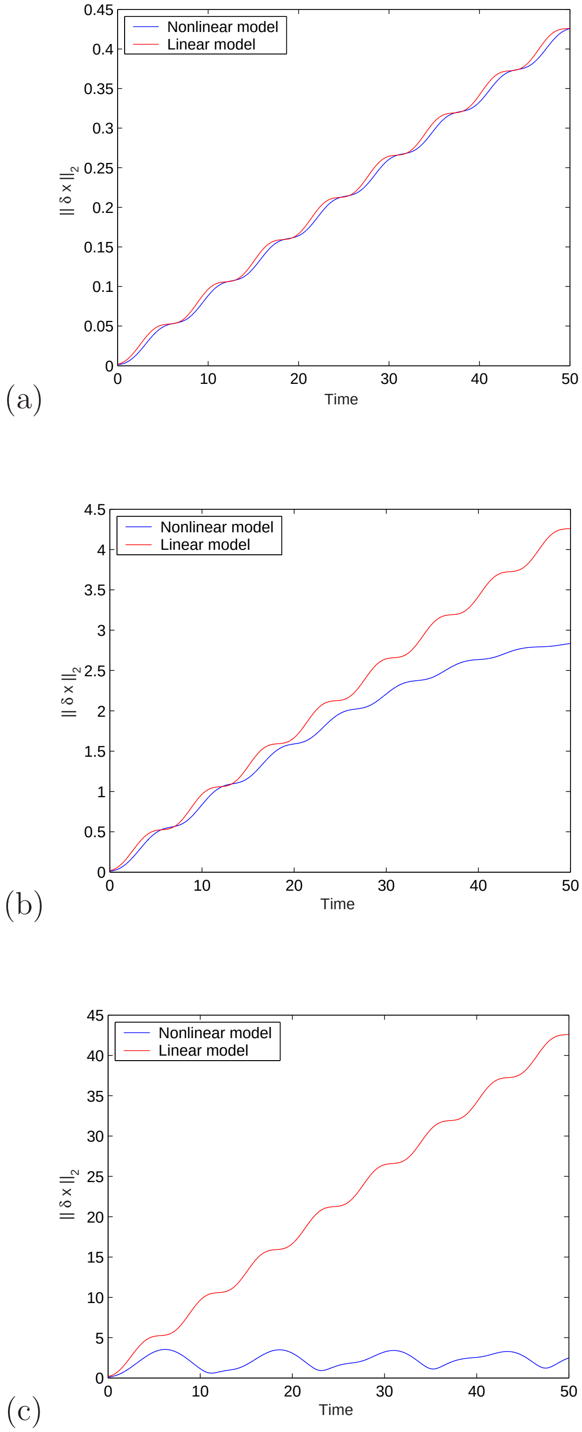


Figure 4.8: Evolution of the perturbations given by the nonlinear and linear models, for (a) $\gamma = 10^{-3}$, (b) $\gamma = 10^{-2}$ and (c) $\gamma = 10^{-1}$; $h = 0.001$, $e = 0$

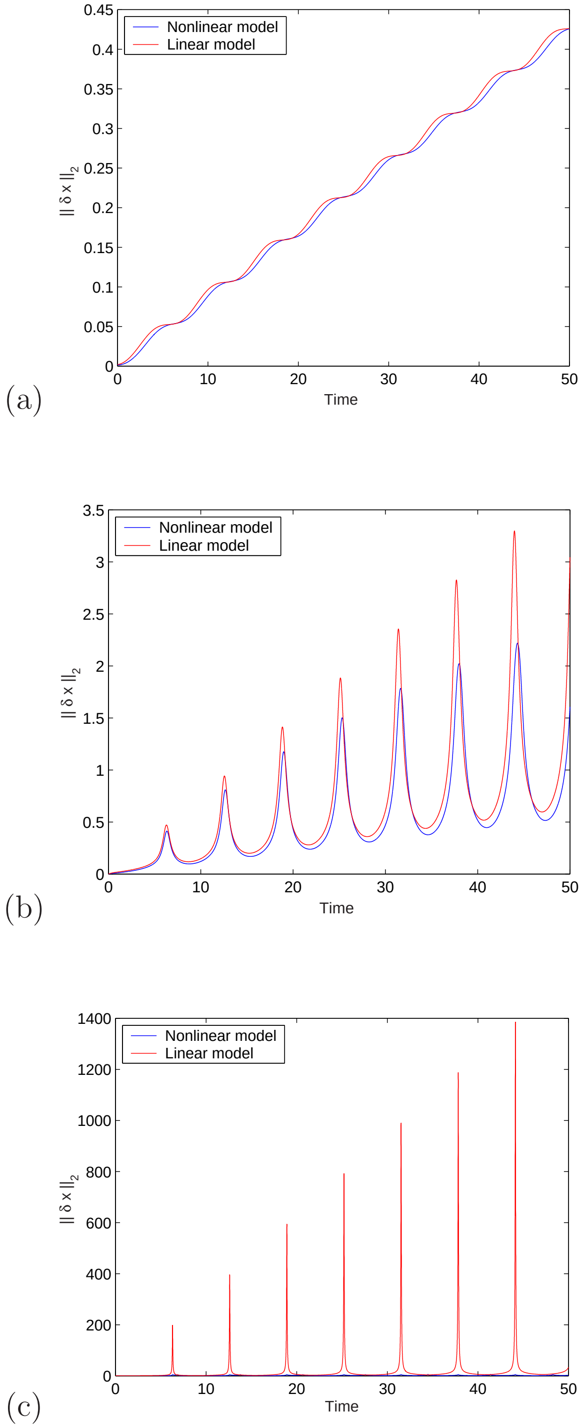


Figure 4.9: Evolution of the perturbations given by the nonlinear and linear models, for (a) $e = 0$, (b) $e = 0.5$ and (c) $e = 0.9$; ($h = 0.001$, $\gamma = 10^{-3}$)

becomes much more extreme as we increase the eccentricity further to $e = 0.9$. Here the linear model departs from the non-linear almost immediately. This qualitative relationship between validity time and eccentricity backs up the results found earlier, giving further evidence that more eccentric ellipses exhibit more non-linear behaviour.

4.3.2 Conservation Properties of the Linear Model

As we have seen, the nonlinear problem conserves the total energy of the system. We now want to consider whether the corresponding linear equations have an analogous conservation property. We consider the linearised equations of the continuous problem, not the discrete system.

$$\frac{\mathrm{d}\delta q_1}{\mathrm{d}t} = \delta p_1 \tag{4.44}$$

$$\frac{\mathrm{d}\delta q_2}{\mathrm{d}t} = \delta p_2 \tag{4.45}$$

$$\frac{\mathrm{d}\delta p_1}{\mathrm{d}t} = -\frac{\delta q_1}{(q_1^2+q_2^2)^{\frac{3}{2}}} + \frac{3q_1(q_1\delta q_1 + q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{5}{2}}} \tag{4.46}$$

$$\frac{\mathrm{d}\delta p_2}{\mathrm{d}t} = -\frac{\delta q_2}{(q_1^2+q_2^2)^{\frac{3}{2}}} + \frac{3q_2(q_1\delta q_1 + q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{5}{2}}}. \tag{4.47}$$

These can be rewritten to give second order equations of the form,

$$\frac{\mathrm{d}^2\delta q_1}{\mathrm{d}t^2} + \frac{\delta q_1}{(q_1^2+q_2^2)^{\frac{3}{2}}} - \frac{3q_1(q_1\delta q_1 + q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{5}{2}}} = 0 \tag{4.48}$$

$$\frac{\mathrm{d}^2\delta q_2}{\mathrm{d}t^2} + \frac{\delta q_2}{(q_1^2+q_2^2)^{\frac{3}{2}}} - \frac{3q_2(q_1\delta q_1 + q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{5}{2}}} = 0. \tag{4.49}$$

To proceed further, we need to also consider the second order non-linear equations,

$$\frac{\mathrm{d}^2 q_1}{\mathrm{d}t} + \frac{q_1}{(q_1^2+q_2^2)^{\frac{3}{2}}} = 0 \tag{4.50}$$

$$\frac{\mathrm{d}^2 q_2}{\mathrm{d}t} + \frac{q_2}{(q_1^2+q_2^2)^{\frac{3}{2}}} = 0. \tag{4.51}$$

We then multiply (4.48) by $\frac{\mathrm{d}q_1}{\mathrm{d}t}$, (4.49) by $\frac{\mathrm{d}q_2}{\mathrm{d}t}$, (4.50) by $\frac{\mathrm{d}\delta q_1}{\mathrm{d}t}$ and (4.51) by $\frac{\mathrm{d}\delta q_2}{\mathrm{d}t}$. We then integrate each with respect to time, and add them together to give

$$\begin{aligned} &\int \tfrac{\mathrm{d}^2\delta q_1}{\mathrm{d}t^2}\tfrac{\mathrm{d}q_1}{\mathrm{d}t}\mathrm{d}t + \int \tfrac{\mathrm{d}^2 q_1}{\mathrm{d}t^2}\tfrac{\mathrm{d}\delta q_1}{\mathrm{d}t}\mathrm{d}t + \int \tfrac{\mathrm{d}^2\delta q_2}{\mathrm{d}t^2}\tfrac{\mathrm{d}q_2}{\mathrm{d}t}\mathrm{d}t + \int \tfrac{\mathrm{d}^2 q_2}{\mathrm{d}t^2}\tfrac{\mathrm{d}q_2}{\mathrm{d}t}\mathrm{d}t \\ &+ \int \tfrac{\delta q_1}{(q_1^2+q_2^2)^{\frac{3}{2}}}\tfrac{\mathrm{d}q_1}{\mathrm{d}t}\mathrm{d}t + \int \tfrac{q_1}{(q_1^2+q_2^2)^{\frac{3}{2}}}\tfrac{\mathrm{d}\delta q_1}{\mathrm{d}t}\mathrm{d}t + \int \tfrac{\delta q_2}{(q_1^2+q_2^2)^{\frac{3}{2}}}\tfrac{\mathrm{d}q_2}{\mathrm{d}t}\mathrm{d}t + \int \tfrac{q_2}{(q_1^2+q_2^2)^{\frac{3}{2}}}\tfrac{\mathrm{d}\delta q_2}{\mathrm{d}t}\mathrm{d}t \\ &- \int \tfrac{3q_1(q_1\delta q_1+q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{5}{2}}}\tfrac{\mathrm{d}q_1}{\mathrm{d}t}\mathrm{d}t - \int \tfrac{3q_2(q_1\delta q_1+q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{5}{2}}}\tfrac{\mathrm{d}q_2}{\mathrm{d}t}\mathrm{d}t = \text{constant} \end{aligned} \tag{4.52}$$

We consider the first two terms together, and we can see that they can be written as

$$\int \frac{\mathrm{d}^2\delta q_1}{\mathrm{d}t^2}\frac{\mathrm{d}q_1}{\mathrm{d}t}\mathrm{d}t + \int \frac{\mathrm{d}^2 q_1}{\mathrm{d}t^2}\frac{\mathrm{d}\delta q_1}{\mathrm{d}t}\mathrm{d}t = \int \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}q_1}{\mathrm{d}t}\frac{\mathrm{d}\delta q_1}{\mathrm{d}t}\right)\mathrm{d}t. \tag{4.53}$$

Evaluating the final integral on the right hand side of (4.53) we have

$$\int \frac{\mathrm{d}^2\delta q_1}{\mathrm{d}t^2}\frac{\mathrm{d}q_1}{\mathrm{d}t}\mathrm{d}t + \int \frac{\mathrm{d}^2 q_1}{\mathrm{d}t^2}\frac{\mathrm{d}\delta q_1}{\mathrm{d}t}\mathrm{d}t = \frac{\mathrm{d}q_1}{\mathrm{d}t}\frac{\mathrm{d}\delta q_1}{\mathrm{d}t} + \text{constant}. \tag{4.54}$$

We follow the same argument for the next two terms in equation (4.52) to give

$$\int \frac{\mathrm{d}^2\delta q_2}{\mathrm{d}t^2}\frac{\mathrm{d}q_2}{\mathrm{d}t}\mathrm{d}t + \int \frac{\mathrm{d}^2 q_2}{\mathrm{d}t^2}\frac{\mathrm{d}\delta q_2}{\mathrm{d}t}\mathrm{d}t = \frac{\mathrm{d}q_2}{\mathrm{d}t}\frac{\mathrm{d}\delta q_2}{\mathrm{d}t} + \text{constant}. \tag{4.55}$$

We must then consider the next four terms of equation (4.52) together. These can be rewritten as

$$\int \frac{1}{(q_1^2+q_2^2)^{\frac{3}{2}}}\left(\delta q_1\frac{\mathrm{d}q_1}{\mathrm{d}t} + q_1\frac{\mathrm{d}\delta q_1}{\mathrm{d}t} + \delta q_2\frac{\mathrm{d}q_2}{\mathrm{d}t} + q_2\frac{\mathrm{d}\delta q_2}{\mathrm{d}t}\right)\mathrm{d}t, \tag{4.56}$$

which in turn can be simplified to

$$\int \frac{1}{(q_1^2+q_2^2)^{\frac{3}{2}}}\frac{\mathrm{d}}{\mathrm{d}t}\left(q_1\delta q_1 + q_2\delta q_2\right)\mathrm{d}t. \tag{4.57}$$

We then evaluate this integral to give

$$\begin{aligned} &\int \tfrac{1}{(q_1^2+q_2^2)^{\frac{3}{2}}}\tfrac{\mathrm{d}}{\mathrm{d}t}\left(q_1\delta q_1 + q_2\delta q_2\right)\mathrm{d}t \\ &= \tfrac{(q_1\delta q_1+q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{3}{2}}} + 3\int \tfrac{q_1\delta q_1+q_2\delta q_2}{(q_1^2+q_2^2)^{\frac{5}{2}}}\left(q_1\tfrac{\mathrm{d}q_1}{\mathrm{d}t} + q_2\tfrac{\mathrm{d}q_2}{\mathrm{d}t}\right)\mathrm{d}t. \end{aligned} \tag{4.58}$$

By using the results found in (4.54), (4.55) and (4.58), we can rewrite (4.52) as

$$\begin{aligned}&\tfrac{\mathrm{d}q_1}{\mathrm{d}t}\tfrac{\mathrm{d}\delta q_1}{\mathrm{d}t}+\tfrac{\mathrm{d}q_2}{\mathrm{d}t}\tfrac{\mathrm{d}\delta q_2}{\mathrm{d}t}+\tfrac{(q_1\delta q_1+q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{3}{2}}}+3\int\tfrac{q_1\delta q_1+q_2\delta q_2}{(q_1^2+q_2^2)^{\frac{5}{2}}}\left(q_1\tfrac{\mathrm{d}q_1}{\mathrm{d}t}+q_2\tfrac{\mathrm{d}q_2}{\mathrm{d}t}\right)\mathrm{d}t\\&-\int\tfrac{3q_1(q_1\delta q_1+q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{5}{2}}}\tfrac{\mathrm{d}q_1}{\mathrm{d}t}\mathrm{d}t-\int\tfrac{3q_2(q_1\delta q_1+q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{5}{2}}}\tfrac{\mathrm{d}q_2}{\mathrm{d}t}\mathrm{d}t=\text{constant}.\end{aligned}\tag{4.59}$$

We can see that in the equation above the fourth term cancels with the final two terms, and so we have our conserved quantity

$$\frac{\mathrm{d}q_1}{\mathrm{d}t}\frac{\mathrm{d}\delta q_1}{\mathrm{d}t}+\frac{\mathrm{d}q_2}{\mathrm{d}t}\frac{\mathrm{d}\delta q_2}{\mathrm{d}t}+\frac{(q_1\delta q_1+q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{3}{2}}}=\text{constant}.\tag{4.60}$$

In terms of our canonical coordinates we thus have

$$p_1\delta p_1+p_2\delta p_2+\frac{(q_1\delta q_1+q_2\delta q_2)}{(q_1^2+q_2^2)^{\frac{3}{2}}}=\text{constant}.\tag{4.61}$$

As we have discussed, the full, nonlinear model conserves the total energy of the system. Considering the conserved quantity given by (4.61) we can see that this is in fact the linearisation of the total energy of the reduced two-body system. Thus the nonlinear model conserves the nonlinear energy and the linearised model conserves the linearised energy. In fact, the linearised model is itself a Hamiltonian system, with the Hamiltonian given by (4.61).

We now test whether this property is conserved by our discrete linear model, found by linearising the discrete nonlinear equations. Figure 4.10 illustrates the evolution of the difference between this quantity given by the model and the truth, using initial conditions given by (4.20) to (4.23). We can see that although the behaviour is periodic, there is no divergence, and the amplitude of the deviation from a constant value is very small, of order 10^{-9}.

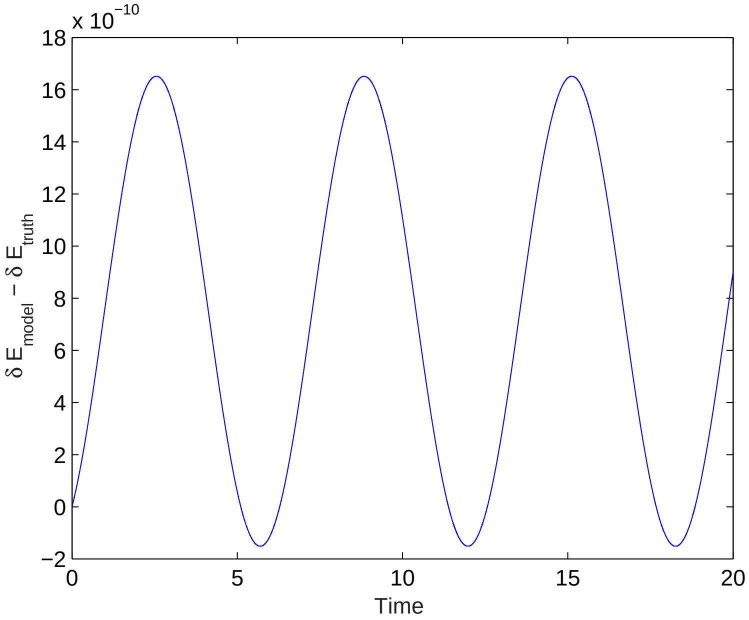


Figure 4.10: Error in the quantity conserved by the linearised two-body model; $e = 0$, $h = 0.001$

4.3.3 The Adjoint Model

In Chapter 3 we discussed the practicalities of the construction of the adjoint model. We have followed the procedure detailed in [17] to create an adjoint for the two-body problem.

A. Adjoint Test

Once the adjoint model is set up, it is useful to show that it is a true adjoint of the cost function. From Section 3.3.3 we have that satisfying

$$< Mx_0, Mx_0 > - < x_0, M^T Mx_0 >= 0, \tag{4.62}$$

where M is the linear model and M^T the adjoint, shows that the code is the true adjoint of the linear model. When this test was carried out for the adjoint model of the two-body system we found that the above was satisfied to machine levels of accuracy, and so we can conclude that the correct adjoint model has been constructed.

B. Gradient Test

We can also test that the adjoint produces the true gradient of the linear model. As described in Section 3.3.3 we must calculate

$$\phi(\alpha) = \frac{J(\boldsymbol{x} + \alpha\frac{\nabla J}{\|\nabla J\|}) - J(\boldsymbol{x})}{\alpha\left(\frac{\nabla J}{\|\nabla J\|}\right)^T, \nabla J} \tag{4.63}$$

for various values of α. Figure 4.11(a) shows the results of plotting $\phi(\alpha)$ versus α, Figure 4.11(b) plots $\log(|\phi(\alpha) - 1|)$ versus α. Figure 4.11(a) should show that for a range of α, $\phi(\alpha) = 1$. These figures compare almost exactly with those illustrated in [30] and thus we can see that our adjoint model does provide the true gradient for the tangent linear model.

4.3.4 The Minimisation Algorithm

To minimise our cost function we use the Conmin minimisation algorithm [46]. As discussed in section 3.3.4 we use the quasi-Newton method for all experiments, with $\varepsilon = 10^{-6}$. Using this value of ε gives the following stopping criteria

$$\frac{|J_{k-1} - J_k|}{1 + |J_k|} < 10^{-6} \tag{4.64}$$

$$\frac{\|\mathbf{x}_{k-1} - \mathbf{x}_k\|}{1 + \|\mathbf{x}_k\|} < 10^{-3} \tag{4.65}$$

$$\frac{\|\nabla J_k\|}{1 + |J_k|} \leq 10^{-2}, \tag{4.66}$$

and all three must be satisfied to stop the minimisation algorithm.

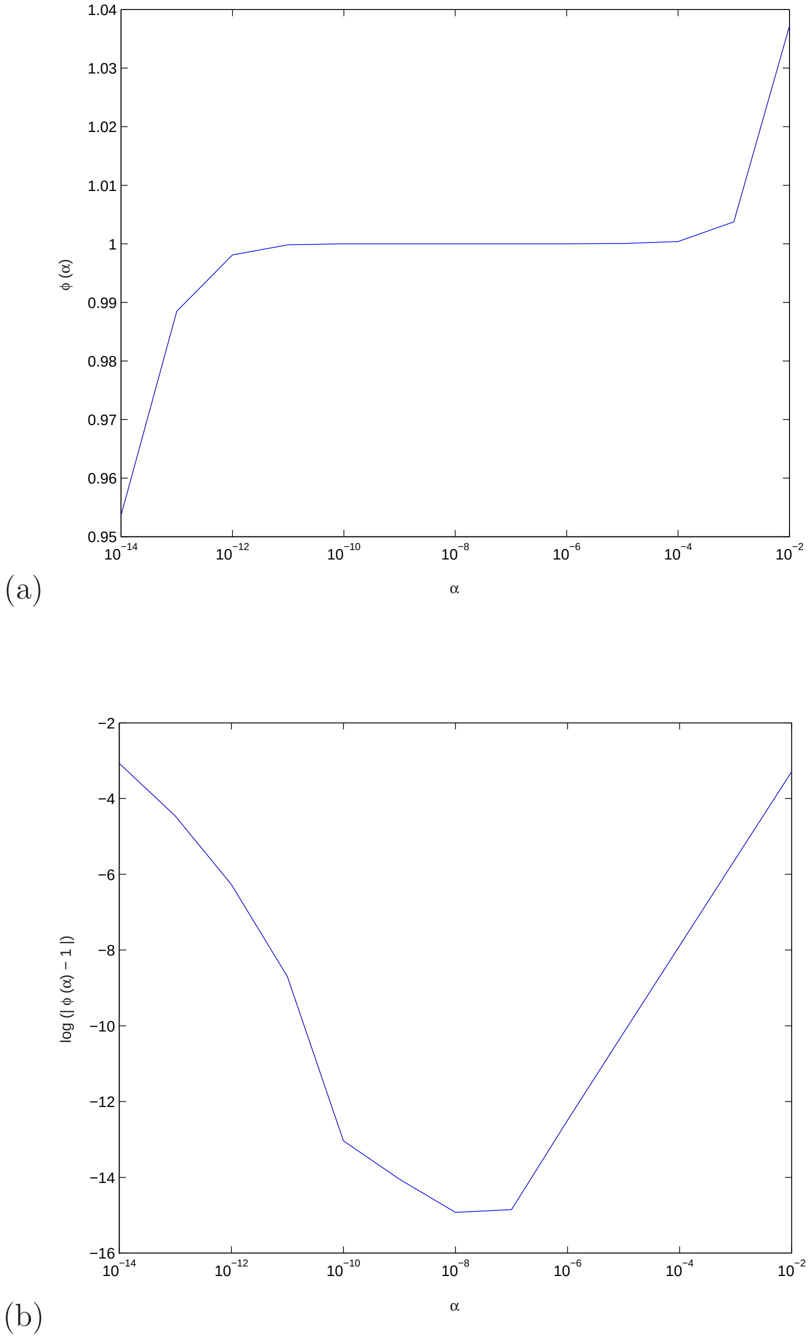


Figure 4.11: Illustration of the gradient tests (a) $\phi(\alpha)$ v α and (b) $\log(|\phi(\alpha) - 1|)$ v $alpha$

4.4 Numerical Experiments

We now carry out a series of experiments using our 4D Var algorithm. For all experiments we consider initial conditions for the true solution to be given by (4.20) to (4.23), with $e = 0$. This choice will avoid any of the difficulties we observed in previous sections relating to highly eccentric orbits and in addition provides a better comparison with the following chapter. As we will see in Chapter 5 the regime we will use for the three-body problem concerns orbits with low eccentricities. In the first instance we include only the observation term in the cost function. For much larger problems we would not be able to do this as the observations alone would not provide enough information and a background term would be needed in order that the problem be well posed. Thus, in the first instance, the cost function we minimise is

$$J(\boldsymbol{x}_0) = \sum_{n=0}^{N} (\boldsymbol{y}_n - H_n[\boldsymbol{x}_n])^T \boldsymbol{R}_n^{-1} (\boldsymbol{y}_n - H_n[\boldsymbol{x}_n]), \tag{4.67}$$

where $\boldsymbol{x}_n = (\boldsymbol{Q}_n, \boldsymbol{P}_n)$, H_n is the observation operator, $\boldsymbol{y}_n$ is the vector of observations and $\boldsymbol{R}_n$ is the observation error covariance matrix. This is subject to the strong constraints that

$$\begin{aligned} \boldsymbol{Q}^n &= \boldsymbol{Q}^{n-1} + h\left(\boldsymbol{P}^{n-1} + \frac{h}{2}\frac{\boldsymbol{Q}^{n-1}}{(Q_1^{n-1^2} + Q_2^{n-1^2})^{\frac{3}{2}}}\right) & (4.68)\\ \boldsymbol{P}^n &= \boldsymbol{P}^{n-1} + \frac{h}{2}\frac{\boldsymbol{Q}^{n-1}}{(Q_1^{n-1^2} + Q_2^{n-1^2})^{\frac{3}{2}}} + \frac{h}{2}\frac{\boldsymbol{Q}^{n}}{(Q_1^{n^2} + Q_2^{n^2})^{\frac{3}{2}}}, & (4.69) \end{aligned}$$

which must be satisfied exactly. These strong constraints are derived from our discrete model equations given by (4.17) to (4.19).

To test our 4D Var scheme we set up an identical-twin experiment. These experiments use the non-linear forward model to provide both the 'true' solution and our observations. Observations that are taken directly from the truth in this way are known as *perfect* observations. We run the 4D Var scheme using the generated observations at every timestep (where $h = 0.001$) to get our optimal initial state. We then run the non-linear model forward from this initial state and compare this trajectory with the truth, using a least squares error. Figure 4.12 illustrates this for the case $e = 0$ and a data assimilation time window $T_{DA} = 12.56$ (approximately two orbits). We can see that the error is very small, of order 10^{-9}, suggesting that our 4D Var scheme is working correctly.

In the last example we used observations taken directly from the true solution. However in general observations have some error associated with them. To account for this error in our system we add noise to the perfect observations. This is done using a random number generator - the resulting noise that is added has a Gaussian distribution and no bias and has a variance of 10^{-4}. We repeat the same

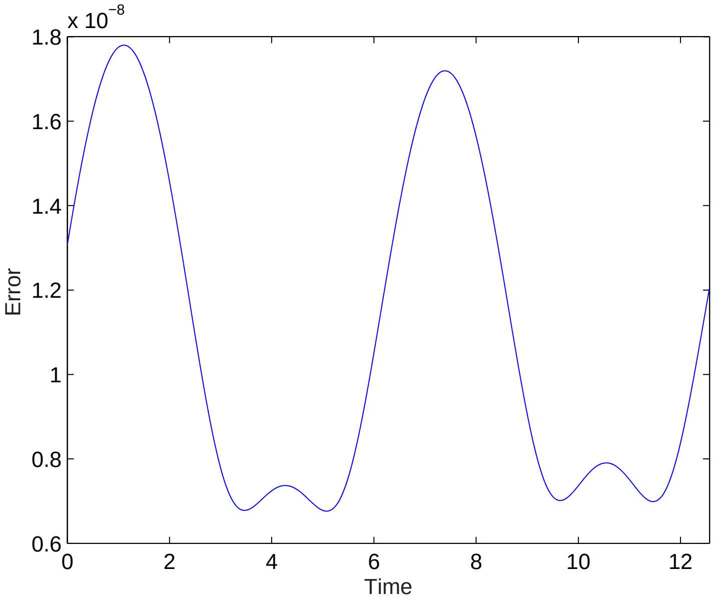


Figure 4.12: Graph to illustrate the error between the optimal solution and the truth using perfect observations at every timestep; $e = 0$, $T_{DA} = 12.56$, $h = 0.001$

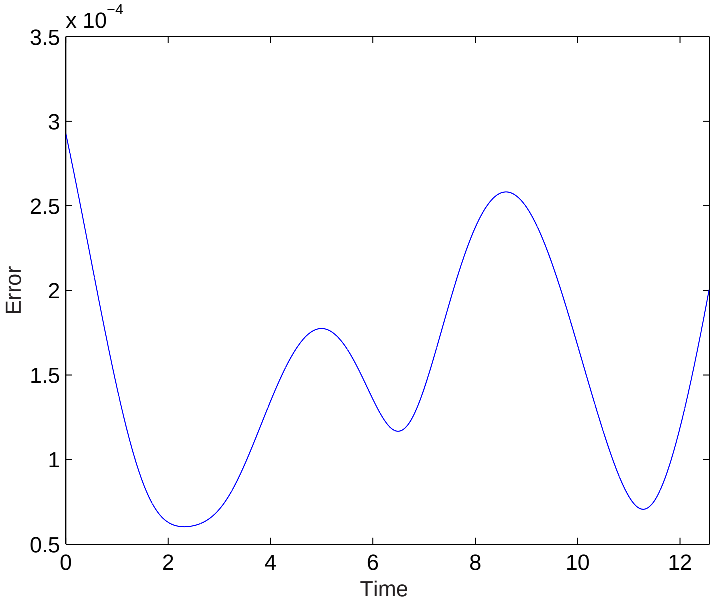


Figure 4.13: Graph to illustrate the error between the optimal solution and the truth using noisy observations at every timestep; $e = 0$, $T_{DA} = 12.56$, $h = 0.001$

identical-twin experiment as before, again using observations in all four variables at each timestep. Figure 4.13 shows the error between the optimal solution and the truth when using these noisy observations. We can see that compared with figure 4.12 the error is much larger, however it is of the same order of magnitude (10^{-4}) as the variance of the observational noise. This is as we would expect, confirming that our scheme is working correctly. From now on all experiments will use noisy observations.

4.4.1 The Effect of Observations

We now consider the effect that our choice of observations has on the analysis of the data assimilation system. For all our future experiments we use a data assmilation window that is half a full orbit. Thus our data assimilation window has length $T_{DA} = \pi$. This is chosen to reflect the system currently in use at the Met Office. There the data assimilation window is six hours corresponding to half an inertial period [12]. This is the approximate period of the gravity waves in the atmosphere.

We consider the effect of using only observations of position *or* momentum. We compare the results using the same number of observations in each case. In our first example we use a dense distribution of observations. In Figure 4.14 we consider the effect of having observations at every timestep of position, (q_1, q_2), only and momentum, (p_1, p_2), only. We compare these with having a full set of observations at every other timestep. Thus the same number of observations are assimilated in each case. Figure 4.14 compares the error in the trajectories of the three cases over the data assimilation window, and a subsequent forecast. We can see that over the data assimilation window the error for each of the three cases is of a similar magnitude. However when we consider the *forecast* we see that, where we have used only observations of position in the 4D Var scheme, the error is increasing more rapidly with time.

A further example that produces this increase in the forecast error is seen when we have sparse observations. In Figure 4.15 we use only eight observations in each case. Where all four components of the state vector are observed these are at $t = \frac{T_o}{4}$ and at $t = \frac{T_o}{2}$, where $T_o = 2\pi$ is the orbital period. Where we assimilate just position or momentum these are included at $t = \frac{T_o}{8}, \frac{T_o}{4}, \frac{3T_o}{8}, \frac{T_o}{2}$. Thus for the case where we use observations of all variables the observation times are more widely spaced. As a result we again observe an increase in the error with time.

In Figures 4.14 and 4.15 we see that within the data assimilation window the analysis has an error that is comparable with the noise on the observations.

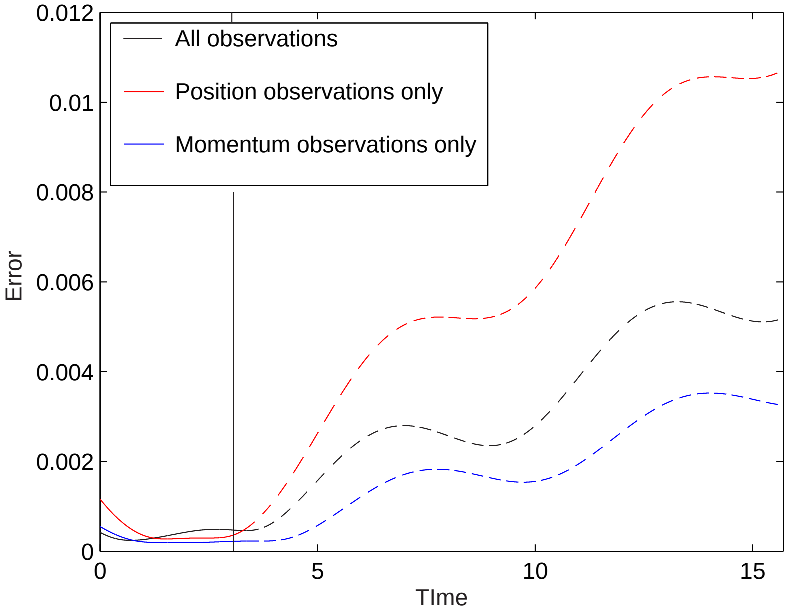


Figure 4.14: Graph to illustrate the error between the optimal solution and the truth using different sets of dense observations; eccentricity $e = 0$, assimilation window $T_{DA} = 3.142$, timestep $h = 0.001$

As we have discussed, this implies that the assimilation algorithm is working. However despite this the errors on the forecasts produced from these analyses are increasing with time. Thus, although we have a good energy conserving model and the data assimilation algorithm is working correctly, the resulting forecasts could be improved. To do this we consider the addition of weak constraints to the cost function.

4.5 Addition of Weak Constraints

In Section 3.4 we have seen that it is possible to impose a constraint that does not have to be exactly satisfied. These are known as weak constraints. In this section

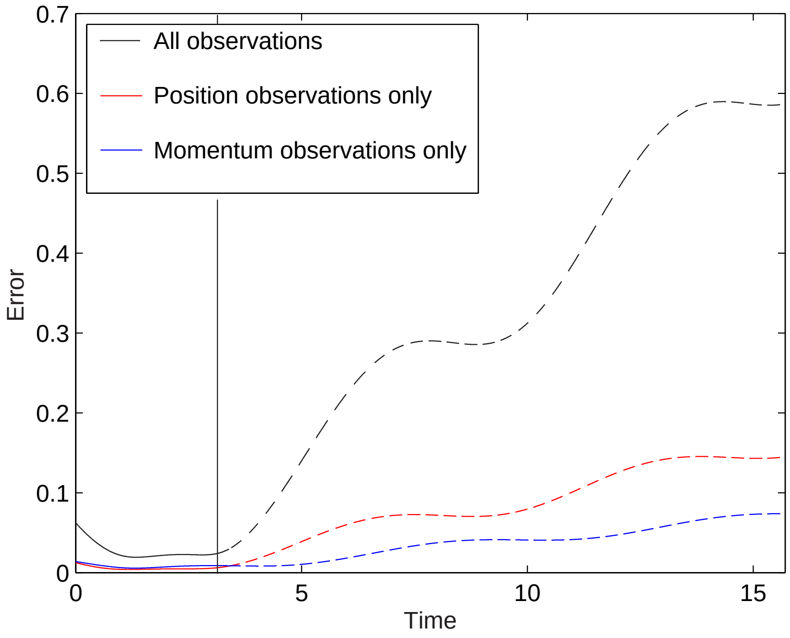


Figure 4.15: Graph to illustrate the error between the optimal solution and the truth using different types of sparse observation, 8 observations only; $e = 0$, assimilation window $T_{DA} = 3.142$

we investigate the addition of different weak constraints, and we test the effect of these on one of the solutions shown in Figure 4.15 where only eight observations were used. We use the example where all of the observations were used at $t = \frac{T_o}{4}$ and at $t = \frac{T_o}{2}$. In these experiments we consider how different constraints can be to used weakly impose that the solutions lie close to a background state.

4.5.1 Perfect Background

In the first instance, in order to compare the types of constraint and understand their effects, we use a perfect background. By this we mean that the background field is equal to the truth. Although this is unrealistic it does allow us to make observations regarding the geometry of the solutions in each case.

The first weak constraint we introduce is equivalent to a typical background term but since there are no background errors the background covariance matrix is equal to the identity matrix. This constraint measures the distance between the model state at the initial time and the background field at the same time. It has the form

$$J_B = \alpha_1(\boldsymbol{x}_b(t_0) - \boldsymbol{x}(t_0))^T(\boldsymbol{x}_b(t_0) - \boldsymbol{x}(t_0)), \tag{4.70}$$

where $\boldsymbol{x}_b = (\boldsymbol{Q}_b, \boldsymbol{P}_b)^T$ is the background state. Here the optimal solution must fit the observations and remain close to the background state vector. Figure 4.16 shows the effect of including this background constraint with $\alpha_1 = 2 \times 10^3, 2 \times 10^4, 2 \times 10^5$, where α_1 is the weighting given to the constraint. We can see that the forecast is considerably better than without any additional constraint.

The second weak constraint makes use of the energy conservation property of the system. Here we measure the distance between the energy of the model state at the initial time and the energy of the background at the same time. It has the form

$$J_E = \alpha_2\left(E(\boldsymbol{x}_b(t_0)) - E(\boldsymbol{x}(t_0))\right)^2 \tag{4.71}$$

In Figure 4.17 we see the effect of including this constraint on the background energy on the same case as for the background constraint. Here we use weightings $\alpha_2 = 2 \times 10^4, 2 \times 10^5, 2 \times 10^6$. These are different values to those assigned to α_1 in Figure 4.16. This is because the magnitude of this term, J_E, is an order of magnitude smaller than J_B and thus we aim to compare terms with an equivalent weight within the cost function. Figure 4.17 shows that the forecast is considerably improved with the inclusion of this term.

We also investigate imposing a weak constraint such that the angular momen-

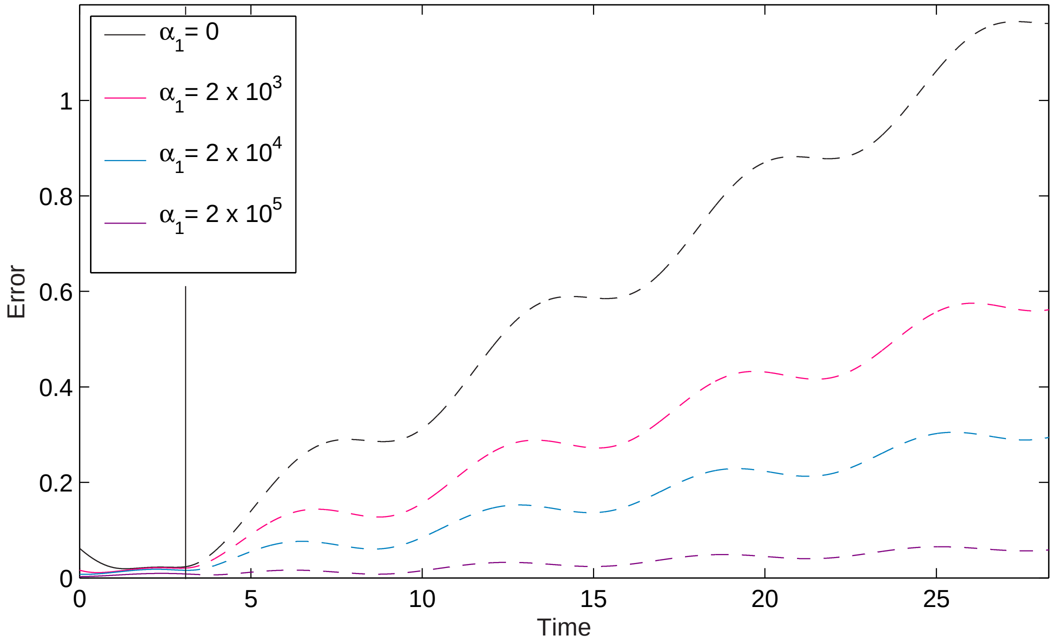


Figure 4.16: The effect of the background constraint on the error in trajectory between optimal solution and truth using sparse observations and a perfect background.

tum of the analysis is close to the angular momentum of the background. This reflects the angular momentum conservation property of the two-body problem discussed in section 4.1.2. This has the form

$$J_L = \alpha_3 \left(L(\boldsymbol{x}_b(t_0)) - L(\boldsymbol{x}(t_0))\right)^2 \tag{4.72}$$

Figure 4.18 illustrates the effect of this constraint on the analysis. Again the background is the true solution and we use values of α_3 equivalent to those of α_2. The results are very similar to those produced using the energy constraint illustrated by Figure 4.17.

We can see that in all cases the constraints improve the forecast. However the behaviour of the solutions is different. In Figure 4.16 we see that the error

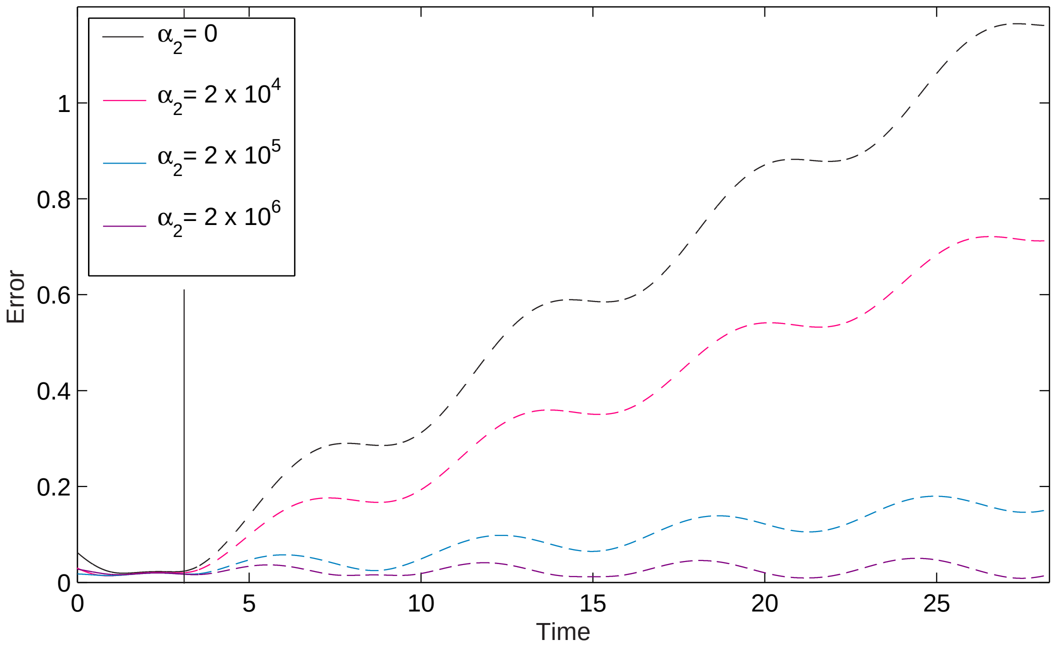


Figure 4.17: The effect of the energy constraint on the error in trajectory between optimal solution and truth using sparse observations and a perfect background.

in the forecast is reduced but it is still increasing and that the amplitude of the oscillations are reduced. By contrast in Figure 4.17 there is also a reduction in the error, however in this case for large values of α_2 the error appears to have stopped increasing. This behaviour is repeated when imposing the angular momentum constraint.

We can explain the difference in the effectiveness of the two constraints by considering the inherent characteristics of the two-body problem. We first consider the effect of the energy constraint. Returning to Figure 4.17 we can see that the error of the unconstrained case is increasing. This behaviour can be caused by comparing two solutions that are out of phase with each other, implying that the orbital periods of the two solutions are different. From Kepler's third law we know

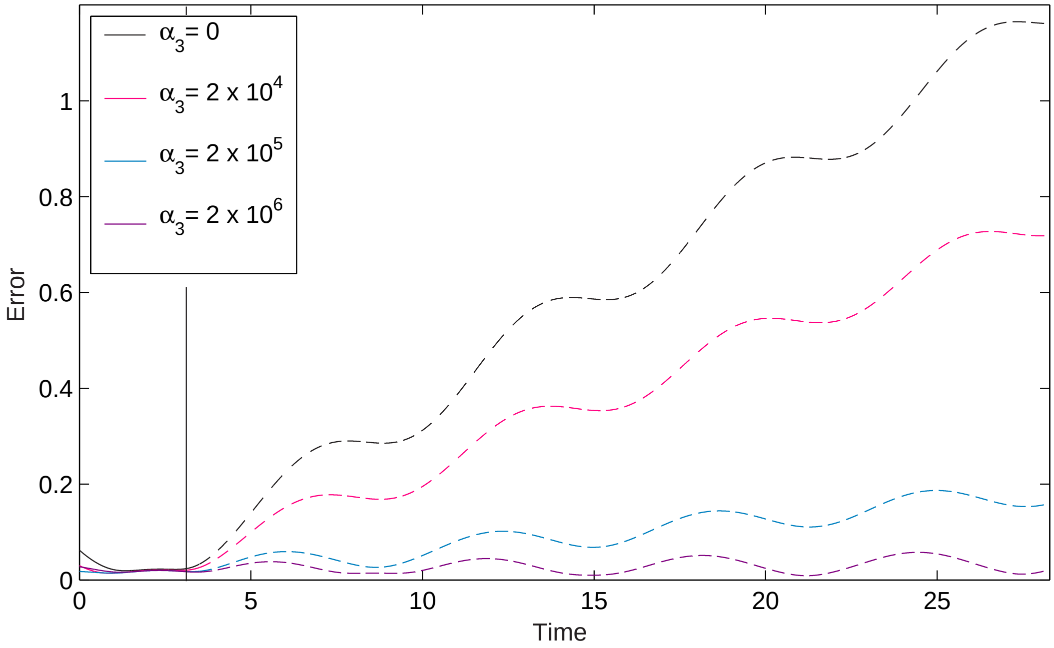


Figure 4.18: The effect of the angular momentum constraint on the error in trajectory between optimal solution and truth using sparse observations and a perfect background.

that a change in the orbital period means a change in the semi-major axis of the orbit. Thus we are comparing two solutions, the analysis and the truth, with different semi-major axes. In addition the energy of a given orbit is only dependent on the semi-major axis [36]. Thus two solutions with different semi-major axes will have different orbital energies.

Thus we can summise that a phase error in the anaysis is due to the analysis having a different energy to the truth. Where we have constrained the energy of the analysis to be close to the energy of the background (which in this example is the truth), this phase error seems to have been eliminated.

We can demonstrate that this increase in error is indeed due to a difference

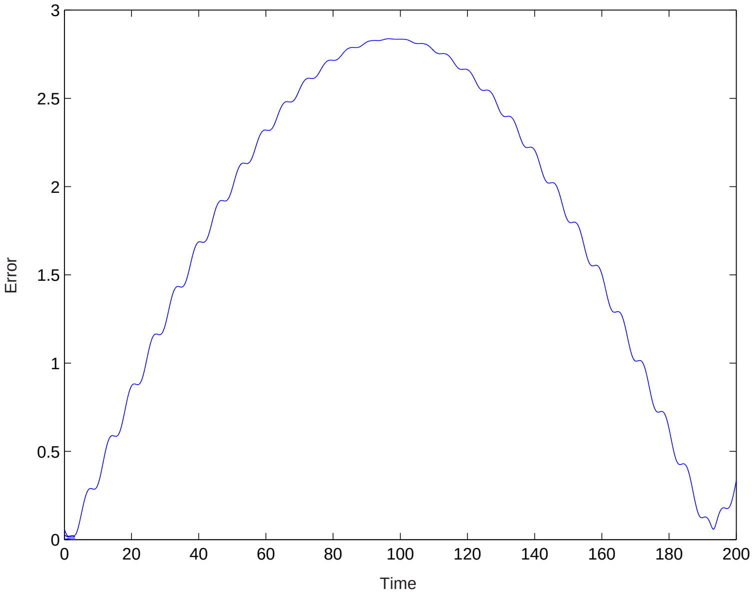


Figure 4.19: Error in trajectory between optimal solution and truth, approx 32 orbits.

in phase by running a long forecast for the unconstrained case. If this is increase is due to a phase difference then the error will increase to a maximum as the two solutions move completely out of phase. This should then begin to decrease as the two orbits move back into phase. Figure 4.19 shows that this is indeed what happens and thus we can conclude that putting a constraint on the energy results in constraining the semi-major axis and period.

We now consider the effect of the background constraint illustrated by Figure 4.16. We recall that in this case the error in the constrained case carried on increasing, however the amplitude of the oscillations was reduced. We can qualitatively explain this by considering the effect of comparing two solutions with different eccentricities but the same semi-major axes. As we have discussed these will have the same period therefore they will always be at the same point on the

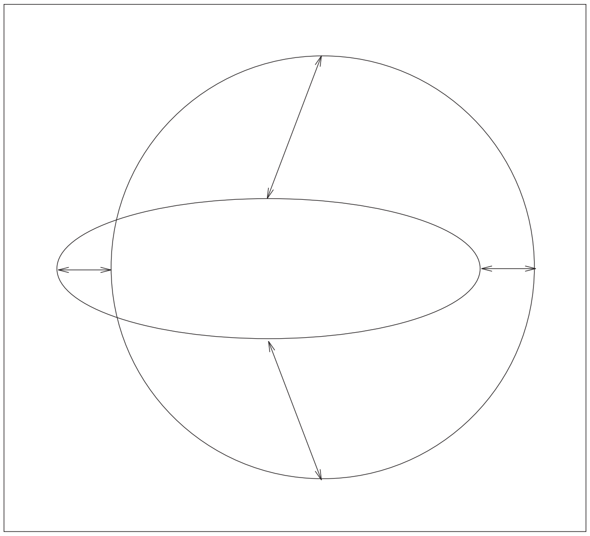


Figure 4.20: Illustration of the effect of comparing two orbits with different eccentricities

orbit, i.e passing the perihelion and aphelion at the same time. This is illustrated by Figure 4.20. Here when the two solutions are compared at the aphelion and perihelion they are closer together than at other points on the orbit. Thus we can see how the oscillations in the error arise.

From this we can see that if the size of the oscillations is reduced then the eccentricities of the two solutions must be closer to the same value. Therefore we can surmise that the background constraint has the effect of constraining the eccentricity of the analysis to be close to the eccentricity of the background.

Finally, as we observed in Figure 4.18, the angular momentum constraint had a very similar effect to the energy constraint. We can thus surmise that this constraint also prevents phase error by constraining the semi-major axis in some way. We recall that, for the two-body problem with closed orbits, Kepler's second law is a consequence of the conservation of angular momentum. This states that a line joining the orbiting body with the focus of the ellipse will sweep out equal areas in equal times. We can therefore infer that a change in a, the semi-major axis, would change this area. Thus by constraining the angular momentum we are constraining the area that is swept out in a given time and thus the semi-major axis.

These experiments illustrate that additional constraints can change the analysis in different ways. The form of the constraint can have a noticeable effect on the geometry of the resulting solution. We have however used an unrealistic set up for these experiments. In reality the background would not be the truth and would have some associated error. In the next section we carry out similar experiments where we will add noise to the background.

4.5.2 Noisy Background

In this section we add the same form of constraints to the cost function as in the previous section. However we now add noise to the background. Again we use the unconstrained example illustrated by figure 4.15 where we assimilate two sets of complete observations at $t = \frac{T_o}{4}, \frac{T_o}{2}$. We then add weak constraints as in the previous section. This time however the background state is no longer perfect. This is achieved by adding random noise with a Gaussian distribution to the truth at the initial time.

The background constraint thus becomes,

$$J_B = (\boldsymbol{x}_b(t_0) - \boldsymbol{x}(t_0))\boldsymbol{B}^{-1}(\boldsymbol{x}_b(t_0) - \boldsymbol{x}(t_0))^T, \tag{4.73}$$

where

$$\boldsymbol{B} = \begin{pmatrix} \sigma_{q_1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{q_2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{p_1}^2 & 0 \\ 0 & 0 & 0 & \sigma_{p_2}^2. \end{pmatrix}$$

Since the values of each of the four variables are of the same magnitude we set the variances, $\sigma_{q_1}^2 = \sigma_{q_2}^2 = \sigma_{p_1}^2 = \sigma_{p_2}^2 = 10^{-4}$.

The energy constraint now has the form

$$J_E = \frac{1}{\sigma_E^2} \left(E(\boldsymbol{x}_b(t_0)) - E(\boldsymbol{x}(t_0)) \right)^2 , \tag{4.74}$$

where the variance σ_E^2 is defined by

$$\sigma_E^2 = \nabla_{\boldsymbol{x}_b} E(\boldsymbol{x}_b(t_0)) \cdot \boldsymbol{\sigma}_{\boldsymbol{x}_b}^2 \tag{4.75}$$

and $\boldsymbol{\sigma}_{\boldsymbol{x}_b}^2 = (\sigma_{q_1}^2, \sigma_{q_2}^2, \sigma_{p_1}^2, \sigma_{p_2}^2)$.

The angular momentum constraint has a similar form to the energy constraint. We add a term to the cost function of the form

$$J_L = \frac{1}{\sigma_L^2} \left(L(\boldsymbol{x}_b(t_0)) - L(\boldsymbol{x}(t_0)) \right)^2 , \tag{4.76}$$

where the variance σ_L^2 is defined by

$$\sigma_L^2 = \nabla_{\boldsymbol{x}_b} L(\boldsymbol{x}_b(t_0)) \cdot \boldsymbol{\sigma}_{\boldsymbol{x}_b}^2 \tag{4.77}$$

and $\boldsymbol{\sigma}_{\boldsymbol{x}_b}^2 = (\sigma_{q_1}^2, \sigma_{q_2}^2, \sigma_{p_1}^2, \sigma_{p_2}^2)$.

We now consider the effect of the addition of these more realistic weak constraints to the original cost function. In Figure 4.21 we illustrate and compare the effect of four different cases,

$$
\begin{aligned}
J &= J_o \\
J &= J_o + J_B \\
J &= J_o + J_E \\
J &= J_o + J_B + J_E.
\end{aligned}
$$

We can see that for this more realistic example the background constraint is considerably more effective than the energy constraint. However when we add both constraints to the cost function this improves the solution even further than with just the background constraint alone.

In Figure 4.22 we consider the angular momentum constraint. We compare the following examples

$$
\begin{aligned}
J &= J_o \\
J &= J_o + J_B \\
J &= J_o + J_L \\
J &= J_o + J_B + J_L.
\end{aligned}
$$

This shows that the angular momentum constraint has almost exactly the same effect as the energy constraint. This backs up the results illustrated by the perfect constraint case. This is demonstrated more clearly in the Figure 4.23. Here

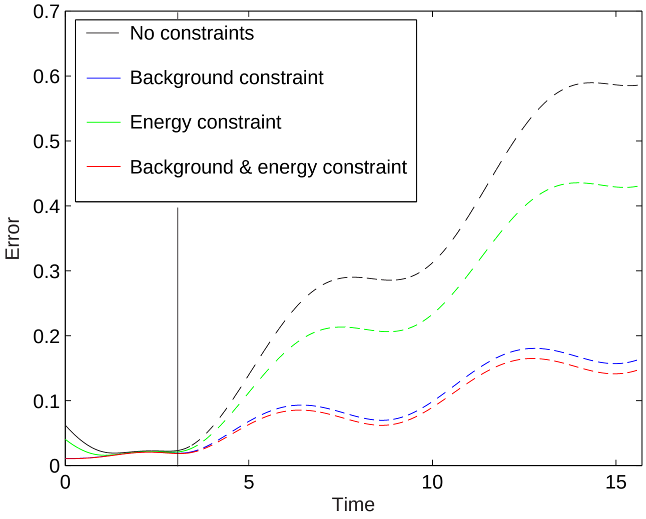


Figure 4.21: The effect of the background and energy constraint on the error in trajectory between optimal solution and truth using sparse observations and a noisy background

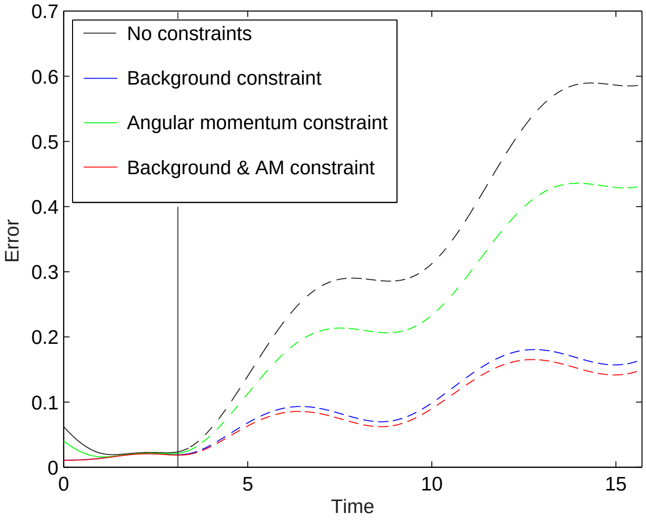


Figure 4.22: The effect of the background and angular momentum constraint on the error in trajectory between optimal solution and truth using sparse observations and a noisy background

we show the effect of the following cases

$$\begin{aligned} J &= J_o \\ J &= J_o + J_B \\ J &= J_o + J_B + J_E \\ J &= J_o + J_B + J_L \\ J &= J_o + J_B + J_E + J_L. \end{aligned}$$

In figure 4.23 we can see more clearly how the angular momentum and energy constraints affect the solution in a very similar manner. However we can also see that if we implement all three of the constraints then the solution is improved even further.

As we have discussed in the previous sections these improvements are a result of constraints being imposed on the geometry of the solution. We have observed that the angular momentum constraint has the same effect as the energy constraint in that it reduces the phase error by constraining the semi-major axis. By adding this constraint to the other two we are not gaining any further information about the structure of the solution. Instead this is equivalent to having twice the weighting on the energy constraint. This is confirmed by Figure 4.24 which compares the results of having all three constraints with the background constraint plus twice the weight of the energy constraint. The results are almost identical.

4.6 Summary and Conclusions

In this chapter we have considered the simple Hamiltonian two-body problem. We derived the continuous canonical equations and explicitly stated the conservation properties of the system. We then derived the discrete model equations

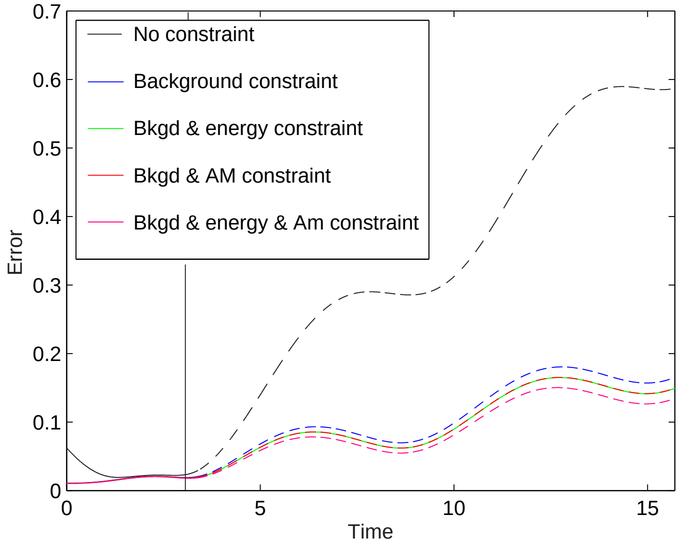


Figure 4.23: The effect of the background, energy and angular momentum constraint on the error in trajectory between optimal solution and truth using sparse observations and a noisy background

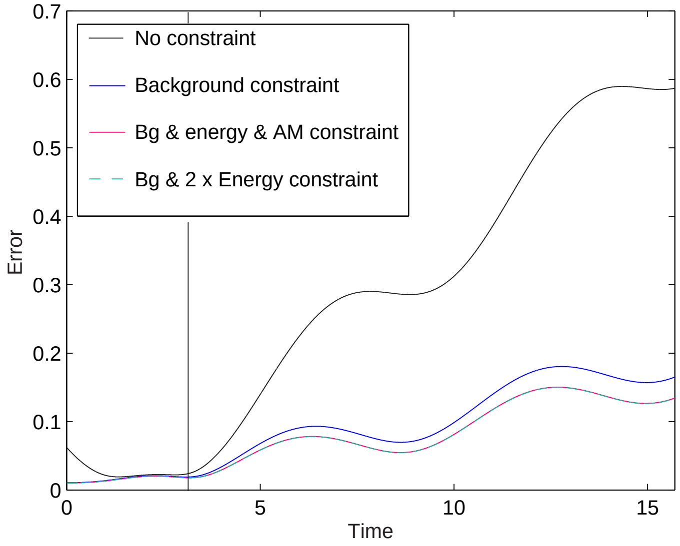


Figure 4.24: The effect of the background, energy and angular momentum constraint on the error in trajectory between optimal solution and truth using sparse observations and a noisy background

using the Störmer-Verlet method and tested the resulting model. We showed that it preserved the inherent characteristics of the original system, specifically the conservation of the Hamiltonian - the total energy of the system.

In Section 4.3 we detailed the steps that were taken in order to construct and test the data assimilation algorithm for this problem. We then went on to describe the numerical experiments that were carried out with the completed 4d Var scheme. We investigated the effect of including additional weak constraints to the system and used these to impose the conservation properties of the system.

We know that our energy conserving nonlinear forward model conserves energy from the initial time. However we wished to impose this as a further constraint such that the analysis would be constrained to lie close to the energy of the background. In practical terms this implies that the constraint would impose this energy conservation property from each data assimilation window to the next. Our results showed that the addition of this energy constraint affected the geometry of the solution. Due to the relationship between energy and the semi-major axis and therefore the period of the orbit we found that the energy constraint reduced the phase error of the solution.

We found this to be true even when the background was noisy. The addition of an energy constraint to a cost function that included an observation term and a background term was shown to give an improved result. We can thus see that the extra term provides further information to the system. In terms of the data assimilation scheme we can explain this by considering the effect of the constraint on the minimisation algorithm. This attempts to find a solution that minimises the observation term and any additional constraint. The gradient of the cost function determines the descent direction of the minimisation algorithm and this must include the gradient of the additional terms. Thus we can infer that the gradient of the cost function will only allow changes to the analysis that satisfy these additional constraints. That is that the analysis must lie close to the background but the change can only be made such that the energy constraint is also satisfied.

In the following chapter we implement a 4d Var scheme for a more complicated Haniltonian system. We will show that the results we have found here can be applied to a more general problem.

Chapter 5

System II: The Three-Body Problem

In Chapter 4 we made use of the information provided in previous chapters to construct a 4d Var assimilation algorithm for the two-body problem. We showed that adding constraints to the cost function that reflected the conservation properties of the system could improve the solution. However the two-body problem is a very simple system and we wish to consider whether these results apply to a more complex Hamiltonian problem.

The full three-body problem has more degrees of freedom than the reduced two-body problem and is a more general system. It can also be set up so that there are two time scales within the problem. This reflects the real problem of the atmosphere where the synoptic flow and the inertial waves have a distinct separation of timescales. As for the two-body system we begin by deriving the canonical equations for the continuous problem and we discuss the conservation properties inherent to this Hamiltonian system. We then must consider our choice of numerical method. Like the two-body system, the three-body problem is an initial value problem and thus the initial conditions will determine the evolution of the system. However the choice of initial conditions is not obvious and thus we will consider several different cases. We then construct and test our 4d Var algorithm in the same way we did in Chapter 4. We finally consider the effect of adding additional constraints to the data assimilation model.

5.1 The Three-Body Problem

In this section we describe the three-body problem and derive both its continuous equations of motion and the discrete equations that will be used in the numerical model. The three-body system describes the evolution of three bodies moving under their mutual gravitational attraction. Each body is therefore affected by the presence of *both* the other bodies. Unlike the two-body problem, this system cannot be solved analytically and it can be shown to exhibit chaotic behaviour.

5.1.1 The Continuous Problem

The three-body system we are using here is that of three bodies with masses m_1, m_2 and m_3, where the motion is confined to a plane. The resulting trajectories of the three bodies are therefore described using vectors of position, $\boldsymbol{q} = (q_1, q_2, q_3, q_4, q_5, q_6)$ and momentum, $\boldsymbol{p} = (p_1, p_2, p_3, p_4, p_5, p_6)$. Thus the vectors $\boldsymbol{q}$ and $\boldsymbol{p}$ contain the position and momentum respectively of all three bodies. The coordinates (q_1, q_2) and (p_1, p_2) refer to the position and momentum of the first body, (q_3, q_4) and (p_3, p_4) to body two and (q_5, q_6) and (p_5, p_6) to the third body.

We simplify the system by using non-dimensional quantities defined by,

$$\begin{aligned}\tilde{m}_i &= \frac{m_i}{M}, && (5.1)\\ \tilde{\boldsymbol{q}} &= \frac{\boldsymbol{q}}{d}, && (5.2)\\ \tilde{t} &= \left(\frac{GM}{d^3}\right)^{\frac{1}{2}} t, && (5.3)\end{aligned}$$

where $\tilde{m}_i$ are the non-dimensional masses, $\tilde{\boldsymbol{q}}$ the non-dimensional position, and $\tilde{t}$ the non-dimensional time. We also have the non-dimensional momentum, $\tilde{\boldsymbol{p}}$ measured in units of $\left(\frac{G}{Md}\right)^{\frac{1}{2}}$. Here G is the universal Gravitational constant and M and d are constants that can be chosen to scale the problem in a suitable manner. Here we set $M = m_1$ and $d = r_{12}$ where r_{12} is the initial distance between the first and second body. For clarity we will drop the $\sim$ from all future notation, although we will always be referring to the non-dimensional problem.

We recall that the first order canonical equations for a system can be derived if we know its Hamiltonian. For this problem, as for the two-body problem, this is defined by,

$$H = T + V, \tag{5.4}$$

where T is the kinetic energy and V the potential energy of the system. In our non-dimensional system the kinetic energy is given by

$$T = \frac{1}{2}\left(\frac{1}{m_1}\left(p_1^2 + p_2^2\right) + \frac{1}{m_2}\left(p_3^2 + p_4^2\right) + \frac{1}{m_3}\left(p_5^2 + p_6^2\right)\right). \tag{5.5}$$

The potential energy is the sum of the potential energy between each pair of bodies. Thus

$$V = -\frac{m_1 m_2}{\left((q_3 - q_1)^2 + (q_4 - q_2)^2\right)^{\frac{1}{2}}} - \frac{m_2 m_3}{\left((q_5 - q_3)^2 + (q_6 - q_4)^2\right)^{\frac{1}{2}}} - \frac{m_3 m_1}{\left((q_1 - q_5)^2 + (q_2 - q_6)^2\right)^{\frac{1}{2}}}. \tag{5.6}$$

Our Hamiltonian is thus

$$\begin{aligned} H &= \frac{1}{2m_1}\left(p_1^2+p_2^2\right)+\frac{1}{2m_2}\left(p_3^2+p_4^2\right)+\frac{1}{2m_3}\left(p_5^2+p_6^2\right) \qquad (5.7)\\ &\quad -\frac{m_1m_2}{\left((q_3-q_1)^2+(q_4-q_2)^2\right)^{\frac{1}{2}}}-\frac{m_2m_3}{\left((q_5-q_3)^2+(q_6-q_4)^2\right)^{\frac{1}{2}}}\\ &\quad -\frac{m_3m_1}{\left((q_1-q_5)^2+(q_2-q_6)^2\right)^{\frac{1}{2}}} \end{aligned}$$

Using (2.25) and (2.26) we can derive the first order, canonical equations for the full three-body problem in a plane

$$\dot{q}_1 = \frac{p_1}{m_1} \qquad (5.8)$$

$$\dot{q}_2 = \frac{p_2}{m_1} \qquad (5.9)$$

$$\dot{q}_3 = \frac{p_3}{m_2} \qquad (5.10)$$

$$\dot{q}_4 = \frac{p_4}{m_2} \qquad (5.11)$$

$$\dot{q}_5 = \frac{p_5}{m_3} \qquad (5.12)$$

$$\dot{q}_6 = \frac{p_6}{m_3} \qquad (5.13)$$

$$\dot{p}_1 = \frac{m_1m_2(q_3-q_1)}{\left((q_3-q_1)^2+(q_4-q_2)^2\right)^{\frac{3}{2}}}-\frac{m_3m_1(q_1-q_5)}{\left((q_1-q_5)^2+(q_2-q_6)^2\right)^{\frac{3}{2}}} \qquad (5.14)$$

$$\dot{p}_2 = \frac{m_1m_2(q_4-q_2)}{\left((q_3-q_1)^2+(q_4-q_2)^2\right)^{\frac{3}{2}}}-\frac{m_3m_1(q_2-q_6)}{\left((q_1-q_5)^2+(q_2-q_6)^2\right)^{\frac{3}{2}}} \qquad (5.15)$$

$$\dot{p}_3 = -\frac{m_1m_2(q_3-q_1)}{\left((q_3-q_1)^2+(q_4-q_2)^2\right)^{\frac{3}{2}}}+\frac{m_2m_3(q_5-q_3)}{\left((q_5-q_3)^2+(q_6-q_4)^2\right)^{\frac{3}{2}}} \qquad (5.16)$$

$$\dot{p}_4 = -\frac{m_1m_2(q_4-q_2)}{\left((q_3-q_1)^2+(q_4-q_2)^2\right)^{\frac{3}{2}}}+\frac{m_2m_3(q_6-q_4)}{\left((q_5-q_3)^2+(q_6-q_4)^2\right)^{\frac{3}{2}}} \qquad (5.17)$$

$$\dot{p}_5 = -\frac{m_2m_3(q_5-q_3)}{\left((q_5-q_3)^2+(q_6-q_4)^2\right)^{\frac{3}{2}}}+\frac{m_3m_1(q_1-q_5)}{\left((q_1-q_5)^2+(q_2-q_6)^2\right)^{\frac{3}{2}}} \qquad (5.18)$$

$$\dot{p}_6 = -\frac{m_2m_3(q_6-q_4)}{\left((q_5-q_3)^2+(q_6-q_4)^2\right)^{\frac{3}{2}}}+\frac{m_3m_1(q_2-q_6)}{\left((q_1-q_5)^2+(q_2-q_6)^2\right)^{\frac{3}{2}}}. \qquad (5.19)$$

These equations need to be solved numerically to find the trajectory of each body in the system.

5.1.2 Conservation Properties

We have seen that the three-body problem differs from the two-body system in several ways. For our purposes there is however an important similarity; both models conserve the Hamiltonian of the system and in both cases this is given by the total energy of the system. Thus the three-body system conserves $H(\boldsymbol{q}, \boldsymbol{p}) = E(\boldsymbol{q}, \boldsymbol{p})$, which is defined by (5.7).

5.1.3 The Discrete Problem

As for the two-body problem, we use the Störmer-Verlet scheme to provide our discrete model. We can see from (5.7) that our Hamiltonian, H, is separable such that

$$H(\boldsymbol{p}, \boldsymbol{q}) = H_1(\boldsymbol{p}) + H_2(\boldsymbol{q}). \tag{5.20}$$

where,

$$\begin{aligned} H_1(\boldsymbol{p}) &= \frac{1}{2m_1}\left(p_1^2 + p_2^2\right) + \frac{1}{2m_2}\left(p_3^2 + p_4^2\right) + \frac{1}{2m_3}\left(p_5^2 + p_6^2\right) && (5.21) \\ H_2(\boldsymbol{q}) &= -\frac{m_1 m_2}{\left((q_3 - q_1)^2 + (q_4 - q_2)^2\right)^{\frac{1}{2}}} - \frac{m_2 m_3}{\left((q_5 - q_3)^2 + (q_6 - q_4)^2\right)^{\frac{1}{2}}} && (5.22) \\ &\quad -\frac{m_3 m_1}{\left((q_1 - q_5)^2 + (q_2 - q_6)^2\right)^{\frac{1}{2}}}. \end{aligned}$$

Using (2.39) to (2.41) we can therefore derive our numerical equations

$$P_1^{n+\frac{1}{2}} = P_1^n + \tfrac{h}{2}\left(\frac{m_1 m_2 (Q_3^n - Q_1^n)}{\left((Q_3^n - Q_1^n)^2 + (Q_4^n - Q_2^n)^2\right)^{\frac{3}{2}}} - \frac{m_3 m_1 (Q_1^n - Q_5^n)}{\left((Q_1^n - Q_5^n)^2 + (Q_2^n - Q_6^n)^2\right)^{\frac{3}{2}}}\right) \tag{5.23}$$

$$P_2^{n+\frac{1}{2}} = P_2^n + \tfrac{h}{2}\left(\frac{m_1 m_2 (Q_4^n - Q_2^n)}{\left((Q_3^n - Q_1^n)^2 + (Q_4^n - Q_2^n)^2\right)^{\frac{3}{2}}} - \frac{m_3 m_1 (Q_2^n - Q_6^n)}{\left((Q_1^n - Q_5^n)^2 + (Q_2^n - Q_6^n)^2\right)^{\frac{3}{2}}}\right) \tag{5.24}$$

$$P_3^{n+\frac{1}{2}} = P_3^n + \tfrac{h}{2}\left(-\frac{m_1 m_2 (Q_3^n - Q_1^n)}{\left((Q_3^n - Q_1^n)^2 + (Q_4^n - Q_2^n)^2\right)^{\frac{3}{2}}} + \frac{m_2 m_3 (Q_5^n - Q_3^n)}{\left((Q_5^n - Q_3^n)^2 + (Q_6^n - Q_4^n)^2\right)^{\frac{3}{2}}}\right) \tag{5.25}$$

$$P_4^{n+\frac{1}{2}} = P_4^n + \tfrac{h}{2}\left(-\frac{m_1 m_2 (Q_4^n - Q_2^n)}{\left((Q_3^n - Q_1^n)^2 + (Q_4^n - Q_2^n)^2\right)^{\frac{3}{2}}} + \frac{m_2 m_3 (Q_6^n - Q_4^n)}{\left((Q_5^n - Q_3^n)^2 + (Q_6^n - Q_4^n)^2\right)^{\frac{3}{2}}}\right) \tag{5.26}$$

$$P_5^{n+\frac{1}{2}} = P_5^n + \tfrac{h}{2}\left(-\tfrac{m_2 m_3 (Q_5^n - Q_3^n)}{\left((Q_5^n - Q_3^n)^2 + (Q_6^n - Q_4^n)^2\right)^{\frac{3}{2}}} + \tfrac{m_3 m_1 (Q_1^n - Q_5^n)}{\left((Q_1^n - Q_5^n)^2 + (Q_2^n - Q_6^n)^2\right)^{\frac{3}{2}}}\right) \tag{5.27}$$

$$P_6^{n+\frac{1}{2}} = P_6^n + \tfrac{h}{2}\left(-\tfrac{m_2 m_3 (Q_6^n - Q_4^n)}{\left((Q_5^n - Q_3^n)^2 + (Q_6^n - Q_4^n)^2\right)^{\frac{3}{2}}} + \tfrac{m_3 m_1 (Q_2^n - Q_6^n)}{\left((Q_1^n - Q_5^n)^2 + (Q_2^n - Q_6^n)^2\right)^{\frac{3}{2}}}\right) \tag{5.28}$$

$$Q_1^{n+1} = Q_1^n + \tfrac{h}{m_1} P_1^{n+\frac{1}{2}} \tag{5.29}$$

$$Q_2^{n+1} = Q_2^n + \tfrac{h}{m_1} P_2^{n+\frac{1}{2}} \tag{5.30}$$

$$Q_3^{n+1} = Q_3^n + \tfrac{h}{m_2} P_3^{n+\frac{1}{2}} \tag{5.31}$$

$$Q_4^{n+1} = Q_4^n + \tfrac{h}{m_2} P_4^{n+\frac{1}{2}} \tag{5.32}$$

$$Q_5^{n+1} = Q_5^n + \tfrac{h}{m_3} P_5^{n+\frac{1}{2}} \tag{5.33}$$

$$Q_6^{n+1} = Q_6^n + \tfrac{h}{m_3} P_6^{n+\frac{1}{2}} \tag{5.34}$$

$$\begin{aligned} P_1^{n+1} = P_1^{n+\frac{1}{2}} + \frac{h}{2}(& \frac{m_1 m_2 (Q_3^{n+1} - Q_1^{n+1})}{\big((Q_3^{n+1} - Q_1^{n+1})^2 + (Q_4^{n+1} - Q_2^{n+1})^2\big)^{\frac{3}{2}}} \\ & - \frac{m_3 m_1 (Q_1^{n+1} - Q_5^{n+1})}{\big((Q_1^{n+1} - Q_5^{n+1})^2 + (Q_2^{n+1} - Q_6^{n+1})^2\big)^{\frac{3}{2}}}) \end{aligned} \tag{5.35}$$

$$\begin{aligned} P_2^{n+1} = P_2^{n+\frac{1}{2}} + \frac{h}{2}(& \frac{m_1 m_2 (Q_4^{n+1} - Q_2^{n+1})}{\big((Q_3^{n+1} - Q_1^{n+1})^2 + (Q_4^{n+1} - Q_2^{n+1})^2\big)^{\frac{3}{2}}} \\ & - \frac{m_3 m_1 (Q_2^{n+1} - Q_6^{n+1})}{\big((Q_1^{n+1} - Q_5^{n+1})^2 + (Q_2^{n+1} - Q_6^{n+1})^2\big)^{\frac{3}{2}}}) \end{aligned} \tag{5.36}$$

$$\begin{aligned} P_3^{n+1} = P_3^{n+\frac{1}{2}} + \frac{h}{2}(& -\frac{m_1 m_2 (Q_3^{n+1} - Q_1^{n+1})}{\big((Q_3^{n+1} - Q_1^{n+1})^2 + (Q_4^{n+1} - Q_2^{n+1})^2\big)^{\frac{3}{2}}} \\ & + \frac{m_2 m_3 (Q_5^{n+1} - Q_3^{n+1})}{\big((Q_5^{n+1} - Q_3^{n+1})^2 + (Q_6^{n+1} - Q_4^{n+1})^2\big)^{\frac{3}{2}}}) \end{aligned} \tag{5.37}$$

$$\begin{aligned} P_4^{n+1} = P_4^{n+\frac{1}{2}} + \frac{h}{2}(& -\frac{m_1 m_2 (Q_4^{n+1} - Q_2^{n+1})}{\big((Q_3^{n+1} - Q_1^{n+1})^2 + (Q_4^{n+1} - Q_2^{n+1})^2\big)^{\frac{3}{2}}} \\ & + \frac{m_2 m_3 (Q_6^{n+1} - Q_4^{n+1})}{\big((Q_5^{n+1} - Q_3^{n+1})^2 + (Q_6^{n+1} - Q_4^{n+1})^2\big)^{\frac{3}{2}}}) \end{aligned} \tag{5.38}$$

$$P_5^{n+1} = P_5^{n+\frac{1}{2}} + \frac{h}{2}\Big(-\frac{m_2 m_3 (Q_5^{n+1} - Q_3^{n+1})}{\big((Q_5^{n+1} - Q_3^{n+1})^2 + (Q_6^{n+1} - Q_4^{n+1})^2\big)^{\frac{3}{2}}} + \frac{m_3 m_1 (Q_1^{n+1} - Q_5^{n+1})}{\big((Q_1^{n+1} - Q_5^{n+1})^2 + (Q_2^{n+1} - Q_6^{n+1})^2\big)^{\frac{3}{2}}}\Big) \tag{5.39}$$

$$P_6^{n+1} = P_6^{n+\frac{1}{2}} + \frac{h}{2}\Big(-\frac{m_2 m_3 (Q_6^{n+1} - Q_4^{n+1})}{\big((Q_5^{n+1} - Q_3^{n+1})^2 + (Q_6^{n+1} - Q_4^{n+1})^2\big)^{\frac{3}{2}}} + \frac{m_3 m_1 (Q_2^{n+1} - Q_6^{n+1})}{\big((Q_1^{n+1} - Q_5^{n+1})^2 + (Q_2^{n+1} - Q_6^{n+1})^2\big)^{\frac{3}{2}}}\Big), \tag{5.40}$$

where $\boldsymbol{Q}^n = (Q_1^n, Q_2^n, Q_3^n, Q_4^n, Q_5^n, Q_6^n)$ and $\boldsymbol{P}^n = (P_1^n, P_2^n, P_3^n, P_4^n, P_5^n, P_6^n)$ are the discrete position and momentum at timestep n.

5.2 Modelling the Three-Body Problem

In this section we discuss the difficulties encountered when modelling the three-body problem with particular reference to our choice of initial conditions, and we test whether the numerical scheme is an appropriate choice for our research.

5.2.1 The Initial Conditions

Once the model is constructed we must make a choice of initial conditions. For the simple, reduced two-body problem an analytic solution can be found. Therefore we *know* that using initial conditions given by (4.20) to (4.23) will result in the body following an elliptical trajectory with eccentricity, e, semi-major axis, $a = 1$ and orbital period, $T = 2\pi$.

The three-body problem cannot be solved analytically and therefore we can only determine numerically the trajectory produced by a given set of initial conditions. For an initial test of the model we use the initial conditions given in [1]. Here however we use momentum, not velocity as Acheson does, and therefore our

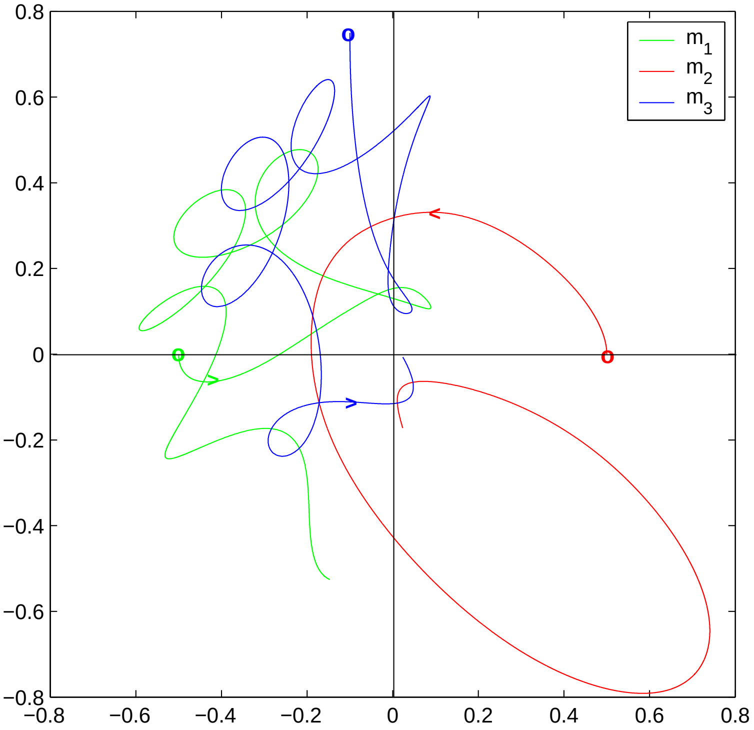


Figure 5.1: Trajectories given by the Störmer-Verlet method with initial positions $(-0.5,\ 0)$, $(0.5,\ 0)$, $(-0.1,\ 0.75)$ and initial momenta $(0,\ -0.15)$, $(0,\ 0.15)$, $(0,\ -0.15)$, $t = 4.85$, timestep $h = 0.001$.

initial conditions have been changed accordingly. The initial conditions are

$$(Q_1, Q_2) = (-0.5, 0) \tag{5.41}$$

$$(Q_3, Q_4) = (0.5, 0) \tag{5.42}$$

$$(Q_5, Q_6) = (-0.1, 0.75) \tag{5.43}$$

$$(P_1, P_2) = (0, -0.15) \tag{5.44}$$

$$(P_3, P_4) = (0, 0.15) \tag{5.45}$$

$$(P_5, P_6) = (0, -0.15), \tag{5.46}$$

and the non-dimensional masses are $m_1 = m_2 = m_3 = 0.5$.

Figure 5.1 shows the trajectories of the three bodies using these initial conditions. Comparing this with Figure 6.15(d) in [1], we can see that the results produced are virtually identical - any small differences are due to a different choice of numerical scheme [1]. From this we can infer our model is working correctly. This figure also illustrates how the three-body problem can behave in an irregular manner. However due to the fact that $m_1 = m_2 = m_3$ we do not have two distinct timescales.

For our research we wish to investigate a three-body system that is predictable and also has two distinct timescales. As we have discussed in Chapter 2 this is to provide an analogy with the atmosphere which also exhibits two different timescales. The initial conditions we took from [1] do not produce a system that behaves in the desired manner. The three-body system that we aim to emulate is that of the Sun, Earth and Moon. We can therefore use real data to try and reproduce such a system. From [41] we have the following

$$r_{SE} = 1.50 \times 10^{11} m \tag{5.47}$$

$$r_{EM} = 3.84 \times 10^{8} m \tag{5.48}$$

$$v_E = 2.98 \times 10^{4} ms^{-1} \tag{5.49}$$

$$v_M = 1.02 \times 10^{3} ms^{-1} \tag{5.50}$$

$$m_S = 1.99 \times 10^{30} kg \tag{5.51}$$

$$m_E = 5.97 \times 10^{24} kg \tag{5.52}$$

$$m_M = 7.35 \times 10^{22} kg \tag{5.53}$$

where r_{SE} is the mean Earth-Sun distance, r_{EM} the mean Earth-Moon distance, v_E an approximation for the orbital velocity of the Earth (using a circular orbit assumption), v_M an approximation for the Moon's orbital velocity and m_S, m_E,

m_M are the masses of the Sun, Earth and Moon respectively. We must first non-dimensionalise these data using (5.1) to (5.3), choosing $M = m_s$ and $d = r_{SE}$. We then convert the velocities to momenta, taking into account that the unit of time is no longer seconds. Our initial conditions based on these data are

$$(Q_1, Q_2) = (0, 0) \tag{5.54}$$

$$(Q_3, Q_4) = (1, 0) \tag{5.55}$$

$$(Q_5, Q_6) = (1, 2.56 \times 10^{-3}) \tag{5.56}$$

$$(P_1, P_2) = (0, 0) \tag{5.57}$$

$$(P_3, P_4) = (0, 3.00 \times 10^{-6}) \tag{5.58}$$

$$(P_5, P_6) = (-1.27 \times 10^{-9}, 3.70 \times 10^{-8}) \tag{5.59}$$

where the masses are $m_1 = 1$, $m_2 = 3 \times 10^{-6}$ and $m_3 = 3.69 \times 10^{-8}$. We have assumed here that the Sun is at the origin and has zero momentum initially. The momentum of the Earth is such that it only has a component in the q_4 direction, the momentum of the Moon in the above initial conditions correponds with zero velocity with respect to the Earth in the q_5 component.

Figure 5.2 shows the trajectory resulting from these initial values. We can see that the system does not behave unpredictably. We also know that it exhibits two separate time scales, the Earth's orbit around the Sun and the Moon's orbit of the Earth. However the trajectories of the Moon and the Earth are very close together compared with the size of the system and we would prefer an example that more obviously illustrates the different paths the bodies follow.

After further investigation we found that the following set of initial conditions provided both different timescales and trajectories that were different enough for

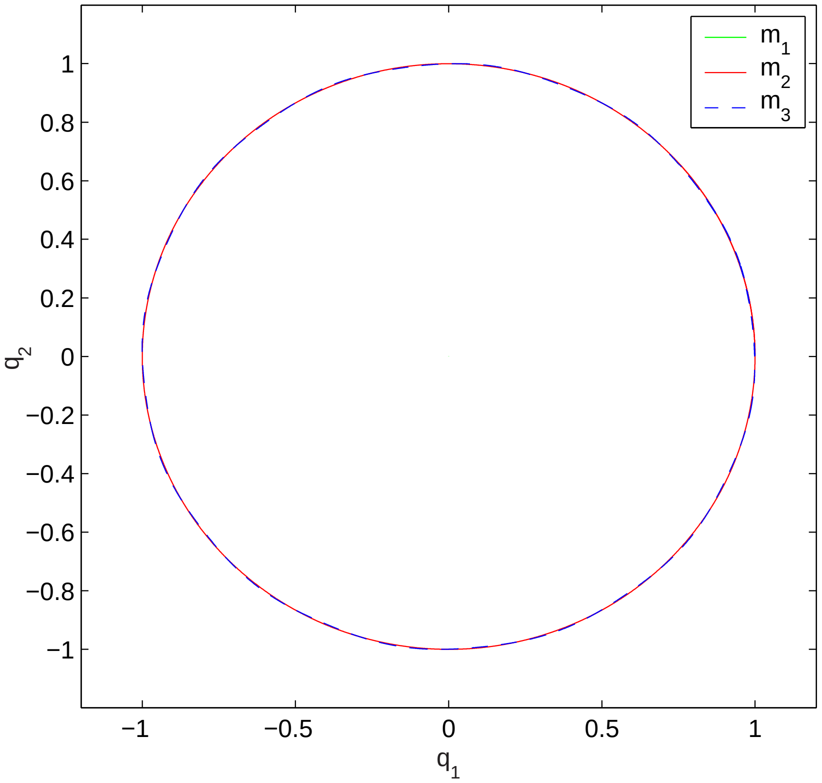


Figure 5.2: Trajectories given by the Störmer-Verlet method with initial positions (0, 0), (1, 0), (1, 2.56×10^{-3}) and initial momenta (0, 0), (0, 3×10^{-6}), (-1.27×10^{-9}, 3.7×10^{-8}), $t = 7.25$, timestep $h = 0.001$.

our purposes.

$$(Q_1, Q_2) = (0, 0) \tag{5.60}$$

$$(Q_3, Q_4) = (1, 0) \tag{5.61}$$

$$(Q_5, Q_6) = (1, 0.1) \tag{5.62}$$

$$(P_1, P_2) = (0, 0) \tag{5.63}$$

$$(P_3, P_4) = (0, 0.1) \tag{5.64}$$

$$(P_5, P_6) = (-0.01, 0.01) \tag{5.65}$$

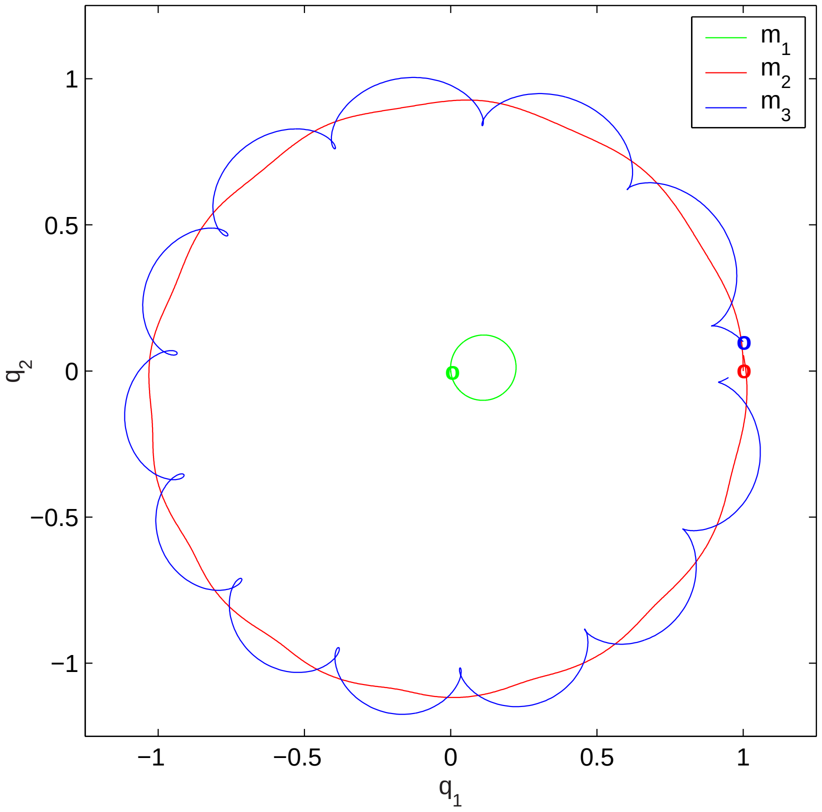


Figure 5.3: Trajectories given by the Störmer-Verlet method with initial positions (0, 0), (1, 0), (1, 0.1) and initial momenta (0.01, −0.11), (0, 0.1), (−0.01, 0.01), $t = 7.25$, timestep $h = 0.001$.

and the masses are $m_1 = 1$, $m_2 = 0.1$ and $m_3 = 0.01$. These can be further improved by choosing the initial momentum for m_1 such that the total linear momentum of the system is zero. This involves changing the initial momentum of m_1; this is now

$$(P_1, P_2) = (0.01, -0.11). \tag{5.66}$$

In Figure 5.3 we see the trajectories of the three bodies from these initial values.

5.2.2 Testing the Model

In the previous section we showed that our model produced the same trajectory as shown in [1] when using the same initial conditions. A further test of the numerical scheme is to consider the total energy of the system. We know that the three-body system is Hamiltonian and thus the total energy of the system is conserved. We therefore look at the energy conserving properties of our model as we did for the two-body problem. Figures 5.4 to 5.6 show the difference between the model energy and the true value for the three different cases shown above.

In the first example, illustrated by Figure 5.4, we have troughs similar to those observed for the two-body problem with a high eccentricity. These corre-

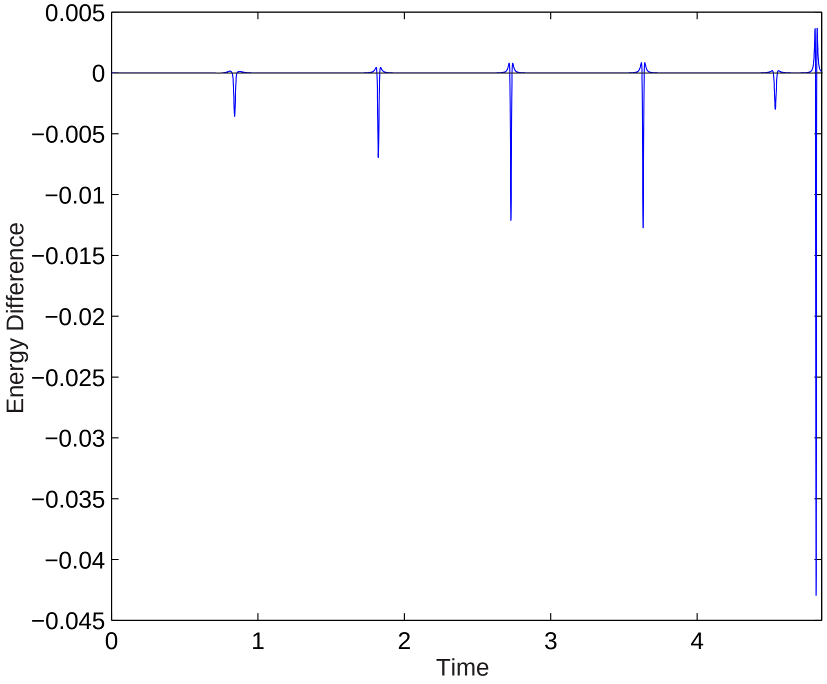


Figure 5.4: Difference between the true energy and the model energy given by the Störmer-Verlet method, initial conditions given by (5.41) to (5.46), $h = 0.001$

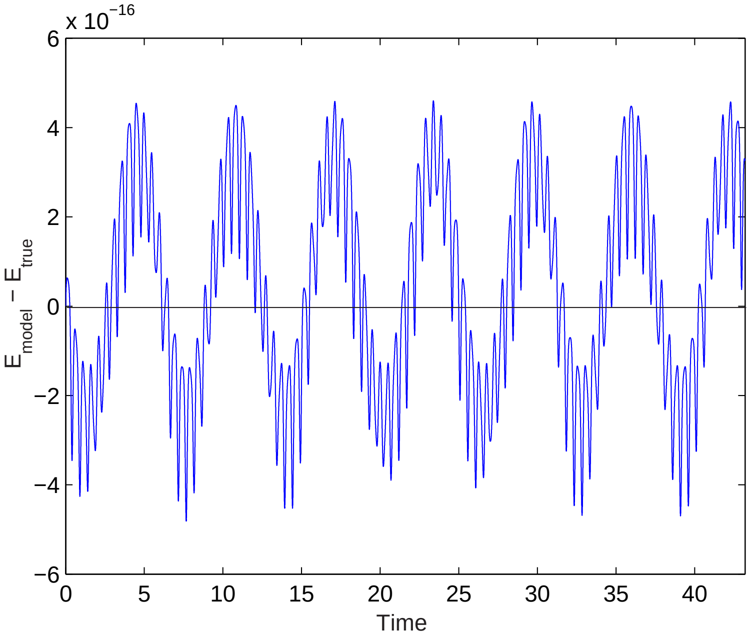


Figure 5.5: Difference between the true energy and the model energy given by the Störmer-Verlet method, initial conditions given by (5.54) to (5.59), $h = 0.001$

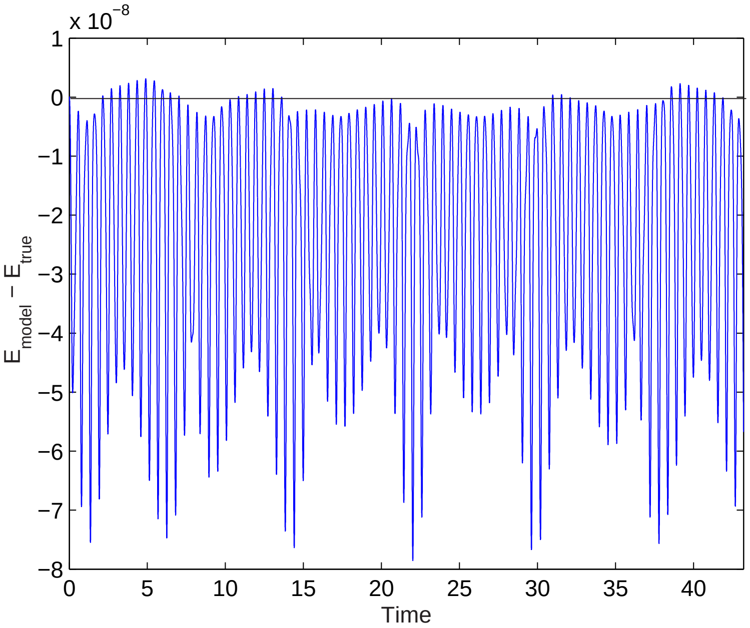


Figure 5.6: Difference between the true energy and the model energy given by the Störmer-Verlet method, initial conditions given by (5.60) to (5.65), $h = 0.001$

spond to the points on the trajectory where a body has a rapid change in direction and thus a large change in velocity due to a close approach with another body. We again see that, although the fixed step method does not do as well in these regions, the energy returns back to its original value after the close approach. For the next two examples, illustrated by Figures 5.5 and 5.6, the difference between the true energy and that given by the model is small - of order 10^{-16} for the first, 10^{-7} for the next. In these two examples the orbits are approximately circular and therefore there are no large changes in orbital velocity that would cause problems with a fixed step scheme.

5.2.3 The Stability of the Three-Body Model

In [49] the stability of the Sun-Earth-Moon system is investigated by considering zero-velocity surfaces. If the orbit of a body lies completely inside a surface of zero velocity then it cannot escape from that orbit and the characteristics of the motion cannot change. This is considered to be a *stable* system. The measure for stability for the general three-body problem is defined in [49] as

$$S = \frac{s_{ac} - s_{cr}}{s_{cr}}. \tag{5.67}$$

If S has a positive value for a particular system then the body's trajectory lies within a zero-velocity surface and thus cannot escape the orbit. In (5.67), s is the stability parameter given by

$$s = -\frac{L^2 H}{G^2 \bar{m}^5}, \tag{5.68}$$

where L is the angular momentum of the system, H is the total energy of the system and $\bar{m}$ is the average mass. The stability parameter s_{ac} is the calculated value for the given problem, s_{cr} is the critical value. For our problem, given by (5.60) to (5.62), (5.66), (5.64) and (5.65), we find that $s_{cr} = 9.057 \times 10^{-2}$. The calculation of S_{cr} is shown in Appendix B. Our problem is non-dimensional therefore the stability parameter becomes

$$s = -\frac{L^2 H}{\bar{m}^5}. \tag{5.69}$$

Total energy, H, and angular momentum, L, are defined [49] as

$$H = -\frac{m_2 m_1}{2a_2}\left[1 + \frac{m_3}{m_2}\left(1 + \frac{m_2}{m_1}\frac{a_2}{a_3}\right)\right] \tag{5.70}$$

$$L = \frac{m_2 m_1}{\sqrt{M}}\sqrt{a_2}\left(\left(1 + \frac{m_1}{m_2}\right)\sqrt{1 - e_2^2} + \frac{m_3}{m_1}\left(\frac{a_3}{a_2}(1 - e_3^2)\right)^{\frac{1}{2}}\left(\frac{M}{m_2 + m_3}\right)^{\frac{1}{2}}\right), \tag{5.71}$$

where a_2, a_3 and e_2, e_3 are the semi-major axes and eccentricities of the orbits of m_2 and m_3, and $M = m_1 + m_2 + m_3$ is the total mass of the system.

In [49] stability is defined by $S > 0$. If S is a *large* positive number then the actual system is not near the critical condition and is thus more stable than a system with a smaller value of S. A system with a negative S is considered to be unstable with an open zero-velocity surface and thus m_3 can escape or enter the orbit of m_2.

We consider the system given by initial conditions (5.60) to (5.62), (5.66), (5.64) and (5.65). We need to estimate the values of the semi-major axes and the eccentricities of the orbit from our numerical trajectories. We find that

$$\begin{aligned} a_2 &= 1.135 \\ a_3 &= 0.09400 \\ e_2 &= 0.09376 \\ e_3 &= 0.1779 \end{aligned}$$

Using these values to calculate H and L we thus find that the stability parameter for this system is $s_{ac} = 8.699 \times 10^{-2}$. Therefore we have that

$$S = \frac{s_{ac} - s_{cr}}{s_{cr}} = -3.952 \times 10^{-2}. \tag{5.72}$$

We can thus see that by this measure the system we are considering is slightly unstable and we can infer that while it may continue to exhibit the same charcteristics of motion, it is possible that m_3 may escape the orbit of m_2 at some point during the evolution.

In Section 5.2 we have investigated a series of different initial conditions for our three-body model. We have shown that the Störmer-Verlet method can reproduce results obtained using other numerical models. We have demonstrated that our numerical scheme provides good conservation of energy and is therefore an appropriate choice for this Hamiltonian problem. We have also shown that our choice of initial conditions produces a slightly unstable configuration. We now consider how sensitive the system is to perturbations of these initial conditions.

5.3 The Sensitivity of the Three-Body Model

The three-body problem differs from the simpler two-body problem in that chaotic behaviour is possible. We have seen in the previous section that the system we are investigating is slightly unstable, but our numerical results, illustrated by Figure 5.3, suggest that the system is stable for the length of the model run. We now wish to consider how sensitive this system is to changes in these initial conditions. We consider how much they can be perturbed without changing the overall behaviour of this system. This is important as, if our data assimilation scheme cannot produce an analysis that is close enough to the truth, then our solution may exhibit completely different behaviour.

We perturb the inital conditions in the following way

$$\tilde{\boldsymbol{Q}}^T = \boldsymbol{Q}^T + \boldsymbol{C}\boldsymbol{Q}^T \tag{5.73}$$

$$\tilde{\boldsymbol{P}}^T = \boldsymbol{P}^T + \boldsymbol{D}\boldsymbol{Q}^T \tag{5.74}$$

where $\boldsymbol{Q} = (Q_1, \ldots, Q_6)$, $\boldsymbol{P} = (P_1, \ldots, P_6)$ and $\boldsymbol{C}$ and $\boldsymbol{D}$ are diagonal matrices with random perturbations of variance σ^2 on the diagonal.

We considered a number of values for σ^2 and found that for values up to $\sigma^2 \approx 0.007$ the behaviour of the three bodies had the same characteristics as the unperturbed problem in that m_3 stayed in orbit around m_2 which itself stayed in orbit around m_1. This is illustrated in figure 5.7.

When the variance is above this value the characteristics of the system change. In Figure 5.8 we show the trajectories for $\sigma^2 = 0.0075$. Here we can see that m_3 is no longer in orbit around m_2, it has switched so that it is in orbit solely around m_1.

In section 5.2.3 we demonstrated that for the unperturbed case the behaviour of the system is slightly unstable. However, in Figure 5.3 we saw that for one large orbit the behaviour given by the numerical model was stable. This can be further illustrated in Figure 5.9 where we run the unperturbed case for 20 orbits. This shows that m_3 will stay in orbit around m_2 which in turn remains in a steady orbit around m_1 for a long model run. Here the orbit of m_3 remains within a bounded region.

However if we run the perturbed case where $\sigma^2 = 0.0075$ for the same period we see that the orbit is less predictable. In Figure 5.10. we see that m_3 mainly orbits m_1 however at one point it is captured by m_2 which it orbits approximately six times before escaping and returning to an orbit around m_1. This corresponds

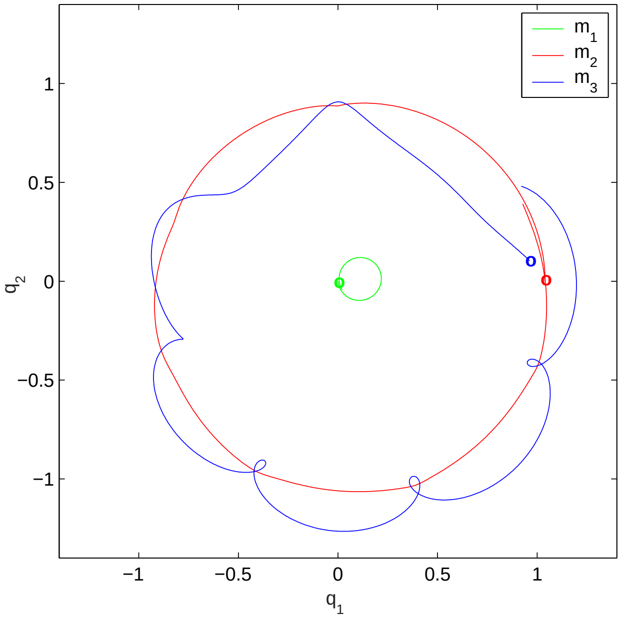


Figure 5.7: Trajectories given by the Störmer-Verlet method with perturbed initial conditions, $\sigma^2 = 0.007$, $t = 7.25$, timestep $h = 0.001$.

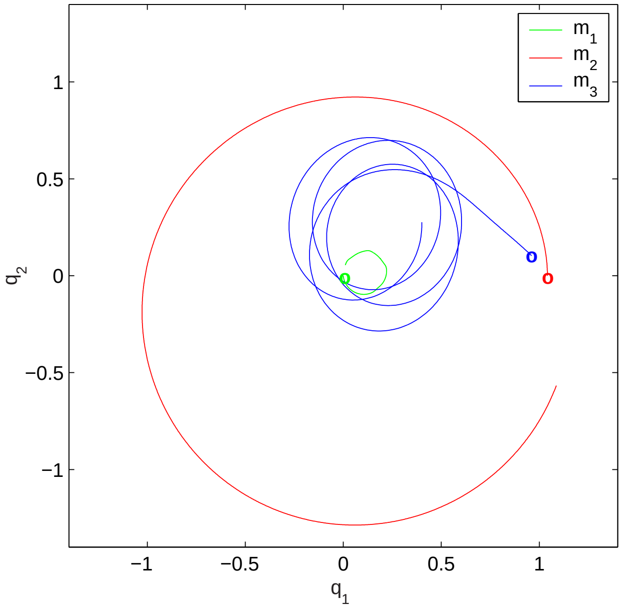


Figure 5.8: Trajectories given by the Störmer-Verlet method with perturbed initial conditions, $\sigma^2 = 0.0075$, $t = 7.25$, timestep $h = 0.001$.

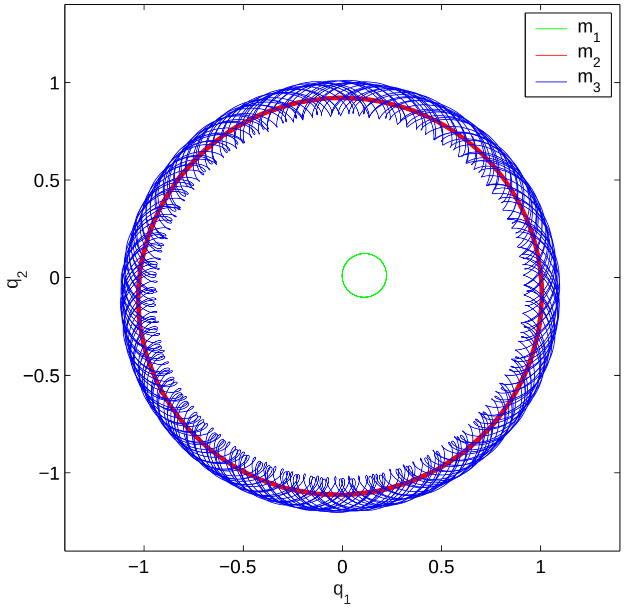


Figure 5.9: Trajectories given by the Störmer-Verlet method with initial positions (0, 0), (1, 0), (1, 0.1) and initial momenta (0.01, −0.11), (0, 0.1), (−0.01, 0.01), $t = 362.5$, timestep $h = 0.001$.

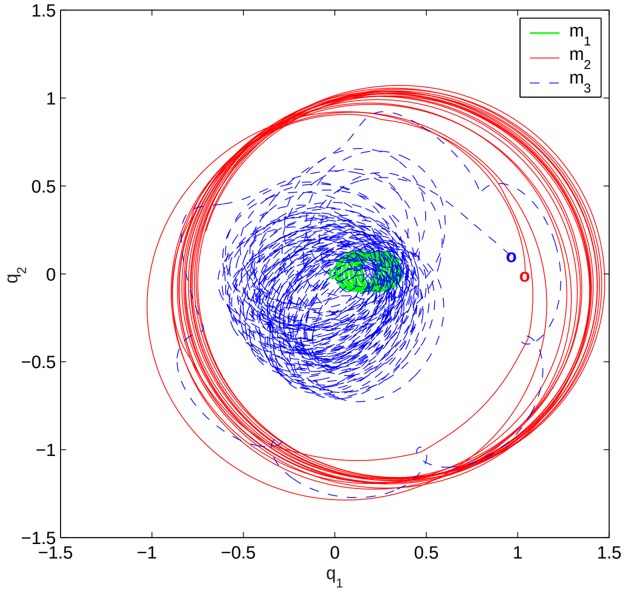


Figure 5.10: Trajectories given by the Störmer-Verlet method with perturbed initial conditions, $\sigma^2 = 0.0075$, $t = 362.5$, timestep $h = 0.001$.

with “very unstable” behaviour as described in [49].

In this section we have demonstrated that, although our system has been shown to be “slightly unstable”, the model behaviour is stable for at least twenty orbits, the length of the model run. This is considerably longer than the time periods we will use in our experiments. We have also shown that small perturbations to these initial conditions can produce completely different behaviour.

5.4 Development of 4D Var for the Three-Body Problem

In Chapter 3 we outlined the stages required for the construction of a 4D Var data assimilation algorithm. In this section we implement each of these for the three-body problem.

5.4.1 The Linear Model

As we did in Section 4.3.1, we follow the procedure outlined in Section 3.2.2 to find our linear model of the three-body problem. By linearising equations (5.23) to (5.40) we create our linearised three-body model equtions. These are listed in Appendix A.

The linear model for the three-body problem must be tested in the same way as the linear two-body problem. As we discussed in Chapter 3 there are two necessary tests.

A. Correctness Test

The first tests the correctness of the linear code. Figure 5.11 plots $||\psi(\delta\boldsymbol{x})||$ versus γ where ψ is defined in (4.43). Again, as for the two-body problem, we see that as $\gamma \to 0$ and therefore $\delta\boldsymbol{x} \to 0$, $||\psi(\delta\boldsymbol{x})|| \to 0$, and thus we can conclude that the linear model is the correct linearisation of our nonlinear three-body problem.

B. Validity Test

We also test the validity of the linear model by following the same procedure as outlined in section 3.3. We run the nonlinear model twice with initial conditions $(\boldsymbol{Q}_0, \boldsymbol{P}_0)$ and then with perturbed initial conditions $((1+\gamma)\boldsymbol{Q}_0, (1+\gamma)\boldsymbol{P}_0)$. We compare the difference between these two non-linear trajectories and the linear model with initial conditions $(\gamma\boldsymbol{Q}_0, \gamma\boldsymbol{P}_0)$. We consider different values of γ and

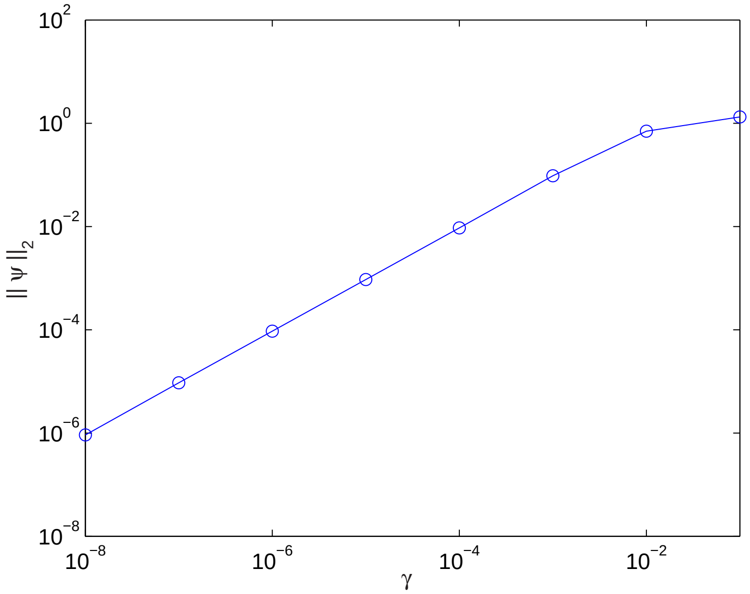


Figure 5.11: Graph to illustrate the correctness of the three-body linear model, initial conditions given by (5.60) to (5.62), (5.66), (5.64) and (5.65), $h = 0.001$.

look at the evolution of the perturbations given by the nonlinear and linear models. Figure 5.12 illustrates the validity times for $\gamma = 10^{-2}$ for each of the three bodies and Figure 5.13 the validity times for $\gamma = 10^{-3}$. Here $\boldsymbol{Q}_0$ and $\boldsymbol{P}_0$ are our final choice of initial conditions given by (5.60) to (5.62), (5.66), (5.64) and (5.65). The red lines represent the nonlinear evolution, the blue lines the linear evolution.

The length of the model run in Figures 5.12 to 5.14 is $t = 7$, corresponding to approximately one orbit of m_2 around m_1. We can see from these plots that for each of the three bodies when $\gamma = 10^{-3}$ the linear model mirrors very closely the behaviour of the nonlinear model. Thus the validity time in this case is relatively long. Where $\gamma = 10^{-2}$ the linear model behaviour is somewhat different to the nonlinear model and this is particularly noticeable for m_3. There is, however, no extreme change in behaviour.

We consider a different choice of initial conditions where $\boldsymbol{Q}_0$ and $\boldsymbol{P}_0$ are given by (5.41) to (5.46) and again look at the validity time. We recall for this case the three masses are equal. Figure 5.14 shows the linear and nonlinear model behaviour for the three bodies where $\gamma = 10^{-3}$. Here we can see that from approximately $t = 6$ the behaviour of the linear model becomes increasingly different to that of the nonlinear model. In Figure 5.15(a) we illustrate the distances between each body, in Figure 5.15(b) we show the momentum of the three bodies for this case. We can see that just before $t = 6$ the distance between m_1 and m_2 is virtually zero. This close approach between the two bodies results in a rapid change in direction for each body. This implies that both will have a rapid change in velocity and therefore momentum, which is illustrated by Figure 5.15(b). As we saw in the two-body problem where there is a close approach, as in the example where eccentricity, $e = 0.9$, the velocity increases as a result of Kepler's second

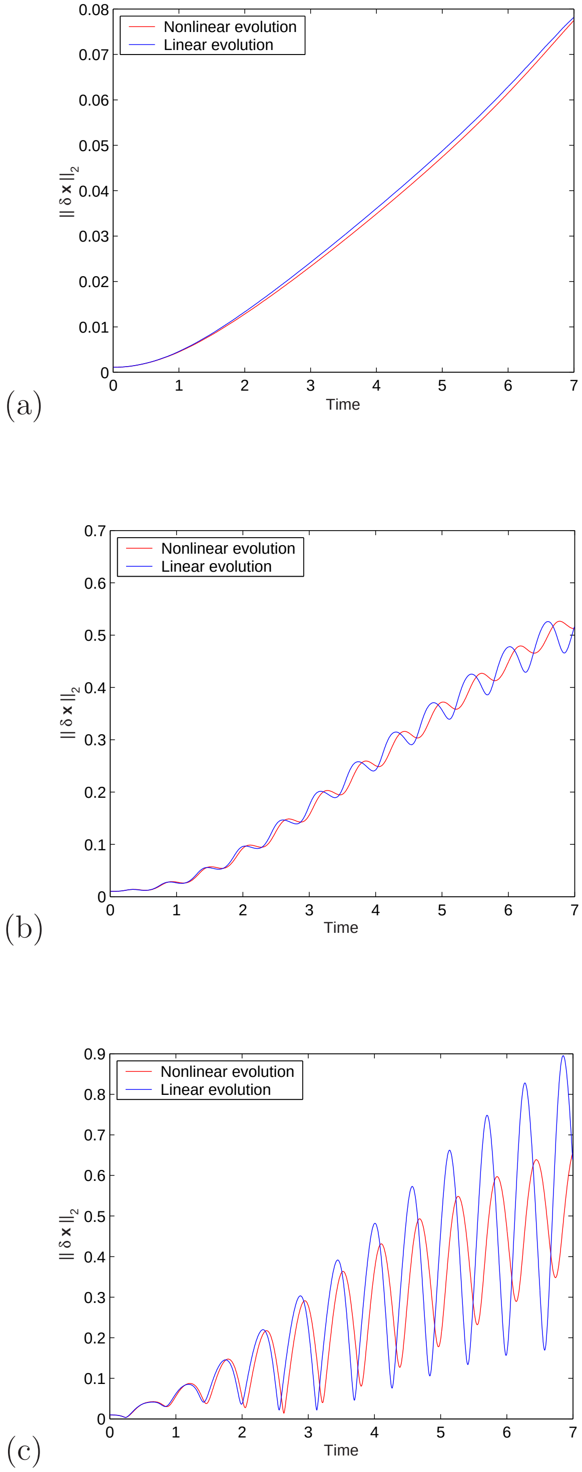


Figure 5.12: Evolution of the position perturbation for the nonlinear and linear models for $\gamma = 10^{-2}$, IC given by (5.60) to (5.62), (5.66), (5.64) and (5.65), (a) body 1, (b) body 2, (c) body 3

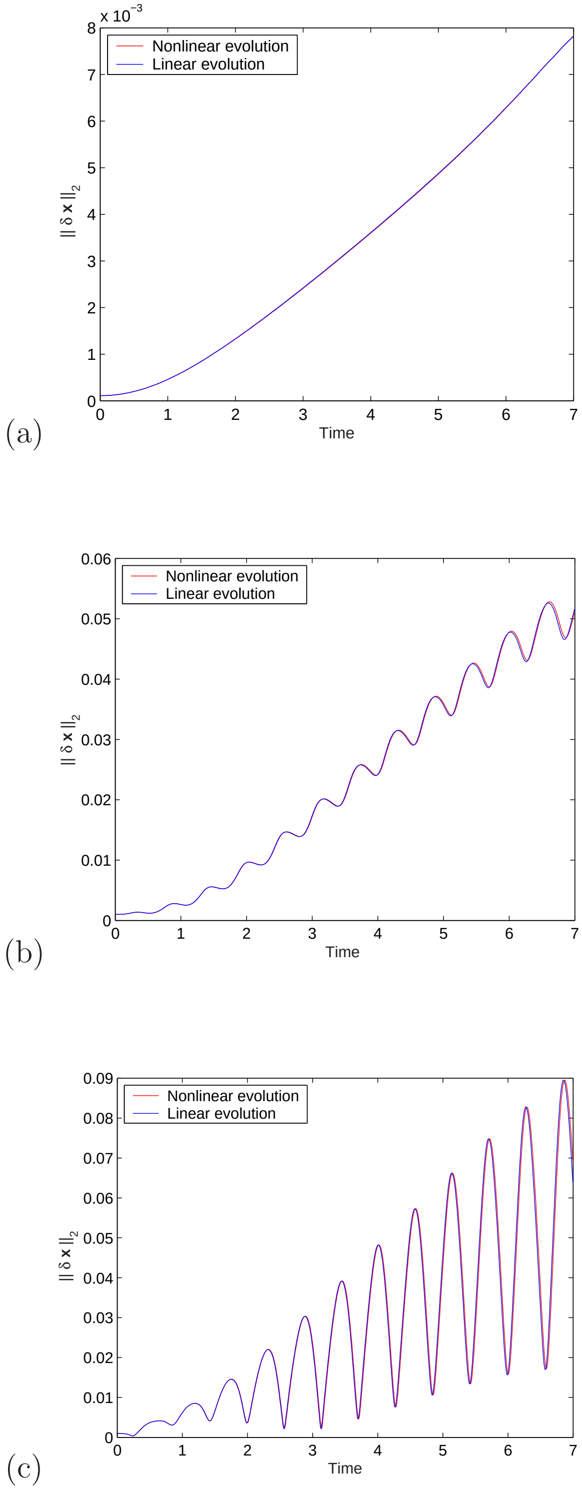


Figure 5.13: Evolution of the position perturbation for the nonlinear and linear models for $\gamma = 10^{-3}$, IC given by (5.60) to (5.62), (5.66), (5.64) and (5.65), (a) body 1, (b) body 2, (c) body 3

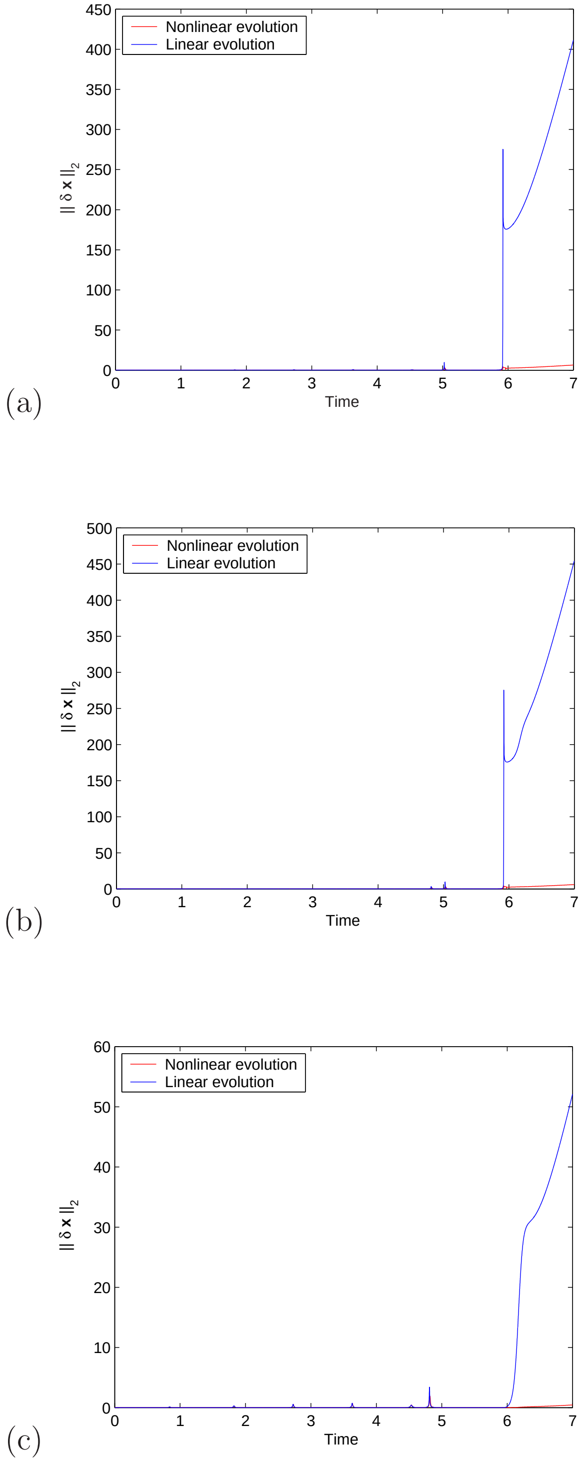
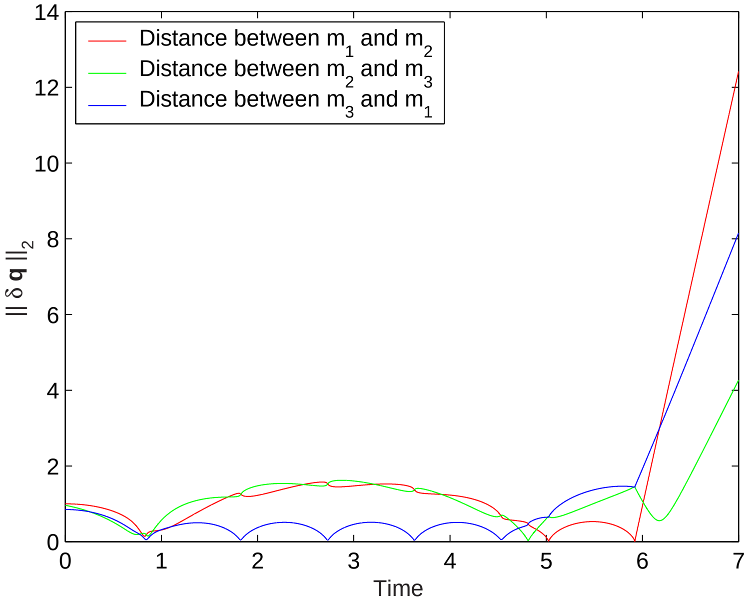
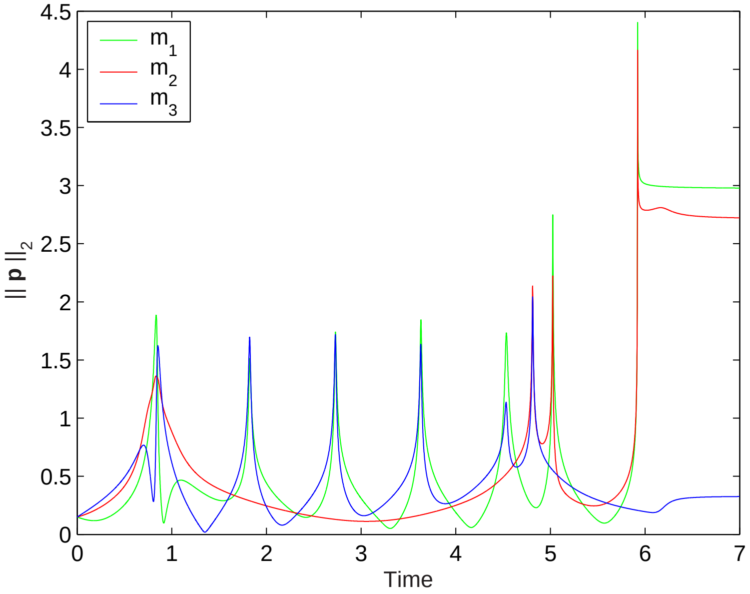


Figure 5.14: Evolution of a perturbation given by the nonlinear and linear models for $\gamma = 10^{-3}$, IC given by (5.41) to (5.46), (a) body 1, (b) body 2, (c) body 3



(a)



(b)

Figure 5.15: Behaviour of (a) the position and (b) the momentum of the system with IC given by (5.41) to (5.46)

law. When such changes occur the linear model is not a valid approximation of the nonlinear model and thus the tangent linear hypothesis does not hold.

From the cases illustrated above we can say that while the three-body problem is in a stable configuration, such as that where m_3 is in orbit around m_2 and they both orbit m_1, the validity time is long and the linear model provides a good approximation for the nonlinear model. In this case the tangent linear hypothesis is valid. Where the bodies become too close the linear model cannot mirror the true behaviour and the tangent linear hypothesis no longer holds.

5.4.2 The Adjoint Model

Again we follow the procedure detailed in [17] to construct our adjoint model directly from the linear model code. We then carry out the two tests required for the adjoint model.

A. Adjoint Test

As discussed in section 3.3.3 the adjoint code is correct if it satisfies

$$< Mx_0, Mx_0 > - < x_0, M^T Mx_0 >= 0, \tag{5.75}$$

where M is the linear model, M^T the adjoint. We find that for the adjoint of the three-body model this is satisfied to machine levels of accuracy.

B. Gradient Test

The previous test shows we have coded the correct adjoint of the tangent linear model but not whether this is the true gradient of the cost function. To do this we plot, in Figure 5.16, $\phi(\alpha)$ versus α and $\log(|\phi(\alpha)-1|)$ versus α where ϕ is defined by

(3.39). We can see that in both cases the figures illustrate the behaviour consistent with the adjoint model being the true gradient [30].

5.4.3 The Minimisation Algorithm

To minimise our cost function we use the Conmin minimisation algorithm [46]. As discussed in section 3.3.4 we use the quasi-Newton method for all experiments with $\varepsilon = 10^{-6}$. Using this value of ε gives the following stopping criteria

$$\frac{|J_{k-1} - J_k|}{1 + |J_k|} < 10^{-6} \tag{5.76}$$

$$\frac{\|\mathbf{x}_{k-1} - \mathbf{x}_k\|}{1 + \|\mathbf{x}_k\|} < 10^{-3} \tag{5.77}$$

$$\frac{\|\nabla J_k\|}{1 + |J_k|} \leq 10^{-2}, \tag{5.78}$$

and all three must be satisfied to stop the minimisation.

5.5 Numerical Experiments

For all our experiments the initial conditions for the true solution will be given by (5.60) to (5.62), (5.66), (5.64) and (5.65). In addition we will use data assimilation window, $T_{DA} = 0.3$. This corresponds to approximately half the period of the small scale motion. This again reflects the window length used by the Met Office where the window is six hours, half the period of the inertial gravity waves - the small scale features of the numerical model for the atmosphere. As for the two-body problem our initial experiments include only the observation term of the cost function. We begin our experiments by testing our 4d Var algorithm for the three-body problem. If our scheme is working correctly then the analysis produced when using perfect observations, taken from the true solution at every timestep, should be very close

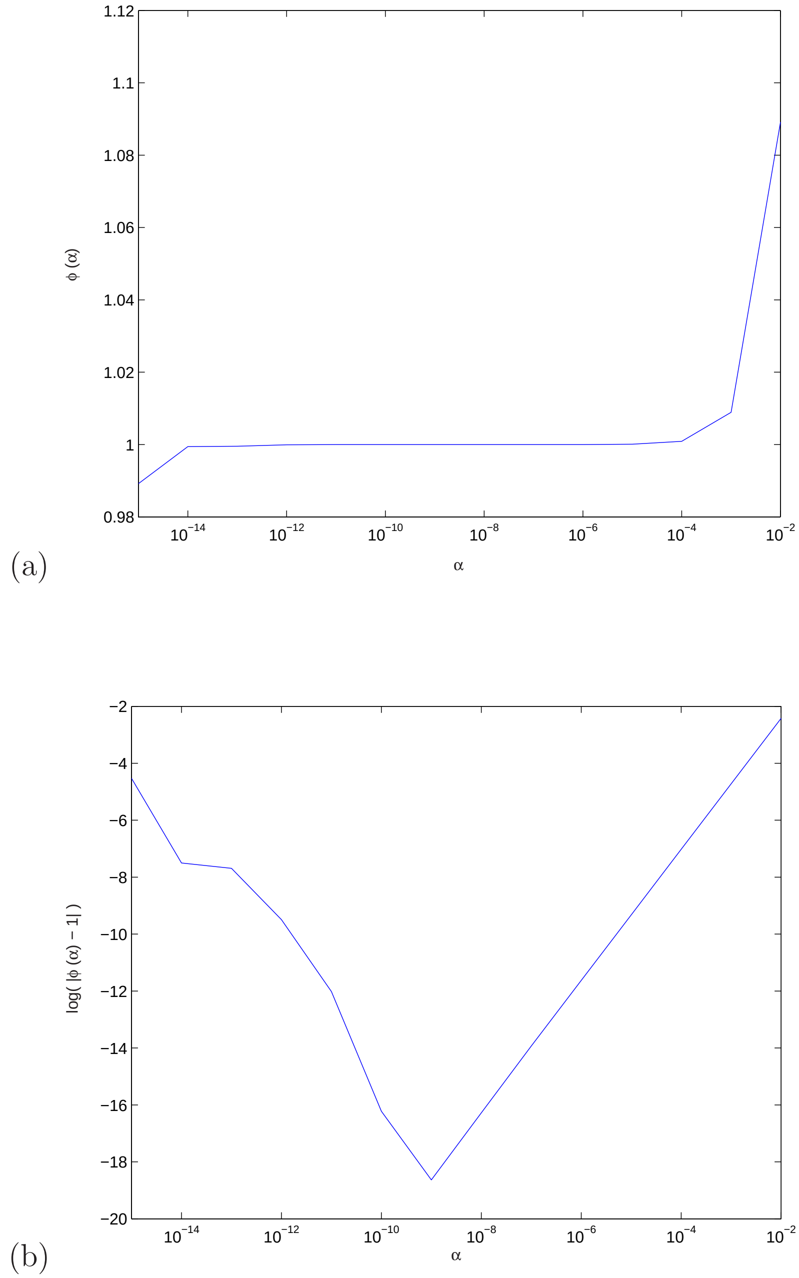


Figure 5.16: Illustration of the gradient tests (a) $\phi(\alpha)$ v α and (b) $\log(|\phi(\alpha) - 1|)$ v $alpha$

to the true solution. Figure 5.17 shows the error in the analysis produced by our scheme for the case of perfect observations.

We can see that the error is of order 10^{-6}. Thus our solution does not lie quite as close to the truth as when we carried this experiment out for the two-body problem. This could be due to the fact that far fewer observations are assimilated here. For the two-body problem we assimilated perfect observations at each timestep with a window of $T_{DA} = 4\pi$ corresponding to two orbital periods. This therefore covers the full trajectory for the orbiting body. Here we are assimilating over a period corresponding to half the period of the small body which is approximately $\frac{1}{24}$th of the large scale orbit. We are therefore assimilating considerably less information relating to the trajectories of the three bodies. We

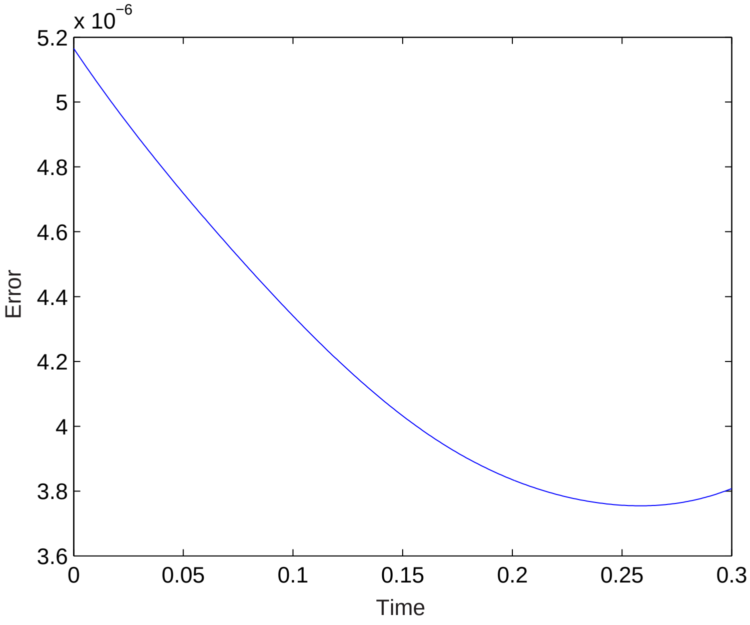


Figure 5.17: Error on the analysis produced by the three-body 4d Var scheme using perfect observations; observations used at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC given by (5.60) to (5.62), (5.66), (5.64) and (5.65)

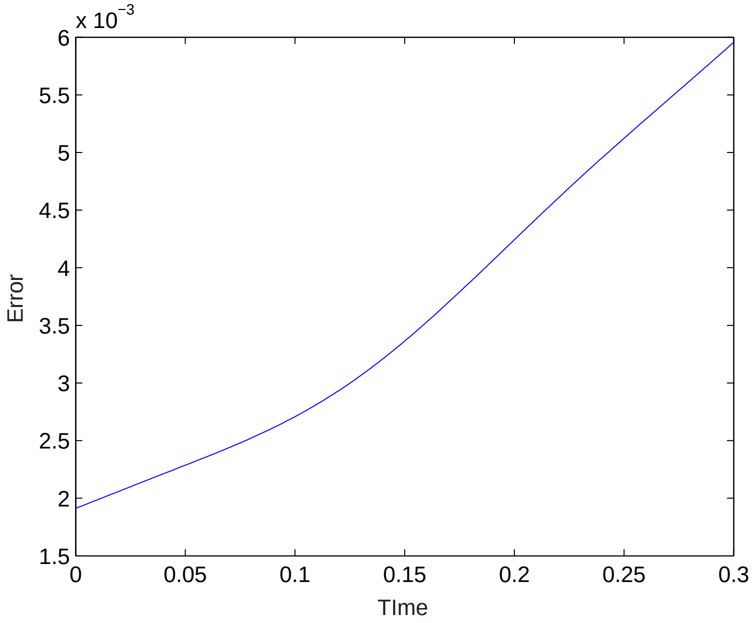


Figure 5.18: Error on the analysis produced by the three-body 4d Var scheme using noisy observations; observations used at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC given by (5.60) to (5.62), (5.66), (5.64) and (5.65)

cannot however test this hypothesis since the minimisation of the 4d Var scheme does not converge for the three-body problem for $T_{DA} > 0.4$. However the error in the analysis is still small and this suggests that our 4d Var scheme is working correctly.

In the next experiment we add noise to our observations to create a more realistic problem. This random noise has a Gaussian distribution with variance that depends on the typical magnitude of the variable. Thus we have $\sigma^2_{Q_1,Q_2,P_1,P_2,P_3,P_4} = 10^{-4}$, $\sigma^2_{Q_3,Q_4,Q_5,Q_6} = 10^{-3}$ and $\sigma^2_{P_5,P_6} = 10^{-5}$. Figure 5.18 shows the error in the analysis for this case. Here we can see that the error is of order 10^{-3}. This confirms that our scheme is working correctly as we would expect the error to be of the order of the largest variance used. For all our remaining experiments we will use noisy observations.

5.5.1 Effect of Observations

In this section we will consider what effect our choice of observations has on the analysis and forecast produced by the data assimilation scheme. In the first of these experiments we use all observations at all timesteps. Figure 5.19 shows the analysis error and the resulting error in the forecast for this case. For all our examples we run the forecast forward for $T_{FC} = 28.8$ which corresponds to approximately four of the large scale orbits. It is necessary to show such a long forecast in order that we can observe both the large scale and the small scale effects of the assimilation.

Here we can see that the error in the smallest body m_3 is the largest. This is as we might expect given that, as we demonstrated in Section 5.3, the smallest body is the most sensitive to changes in the initial conditions.

We now compare this with analyses found using different observation sets. In all cases we use observations at every timestep but we do not use a full set of observations. In the first three examples, illustrated by Figures 5.20 to 5.25, we exclude observations from each of the three bodies in turn. For each case we plot both the error in the forecast and the resulting trajectory.

We can see that in all cases the characteristics of the motion have not changed drastically. The small body, m_3, is still in orbit around m_2 and they are both orbiting m_1. However in Figure 5.20 where we have not assimilated observations for m_1 the magnitude of the error is larger than for the other two cases. In addition we can see from Figure 5.21, we have eleven small orbits for every large orbit compared with twelve in the true solution and the linear momentum is no longer zero.

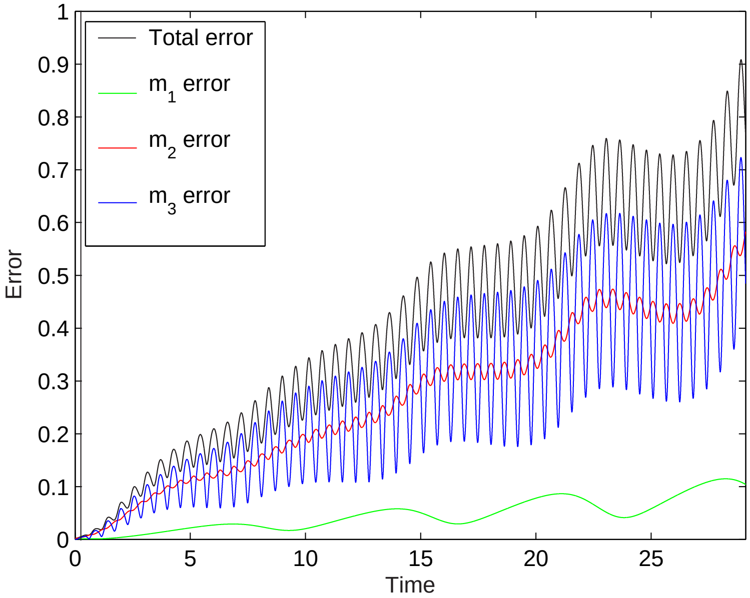


Figure 5.19: Error in the forecast produced by the three-body 4d Var scheme using all observations; observations used at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC given by (5.60) to (5.62), (5.66), (5.64) and (5.65)

In Figures 5.22 and 5.24, where we have no observations of m_2 and m_3 respectively, the magnitude of the error is virtually the same. In Figure 5.23 we have eleven small orbits to one large and there is less transverse motion. In Figure 5.25 we see there are only six small orbits for every large orbit.

From these figures we can conclude that the best solution is found where we have no observations for the the middle body, m_2. We might expect this due to the fact that the observations from m_3 will help to reconstruct both the large and small motion since this has motion on both the scales. In addition observations of m_1 will prevent the entire system from drifting.

In the next examples we use observations from one body only. We note that where we have observations of only the largest body, m_1, or the middle body,

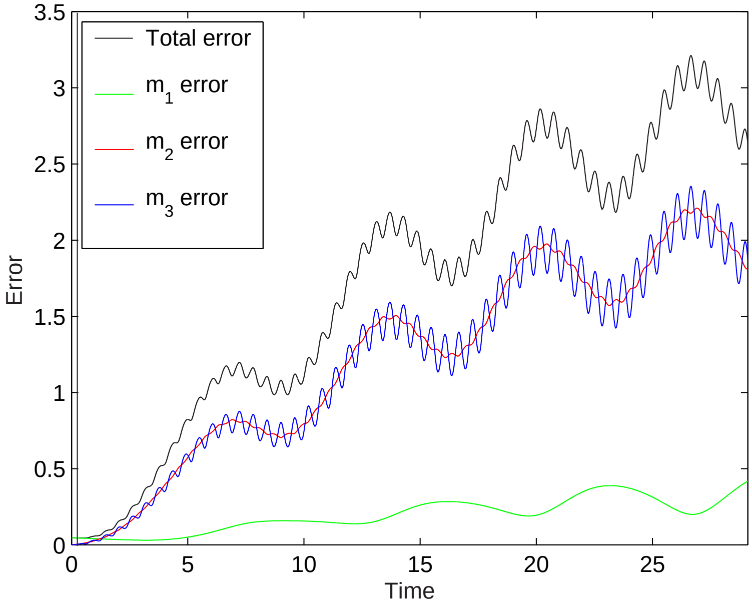


Figure 5.20: Error in the forecast - no observations of m_1; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

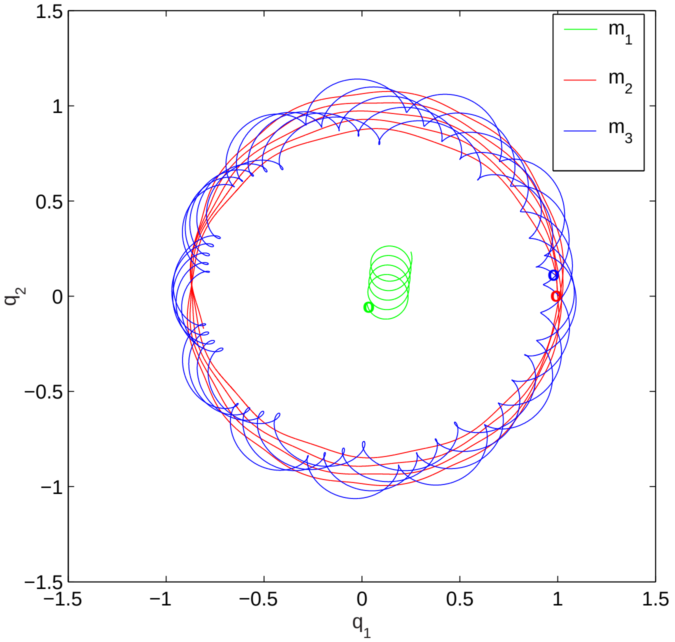


Figure 5.21: Trajectory of the forecast - no observations of m_1; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

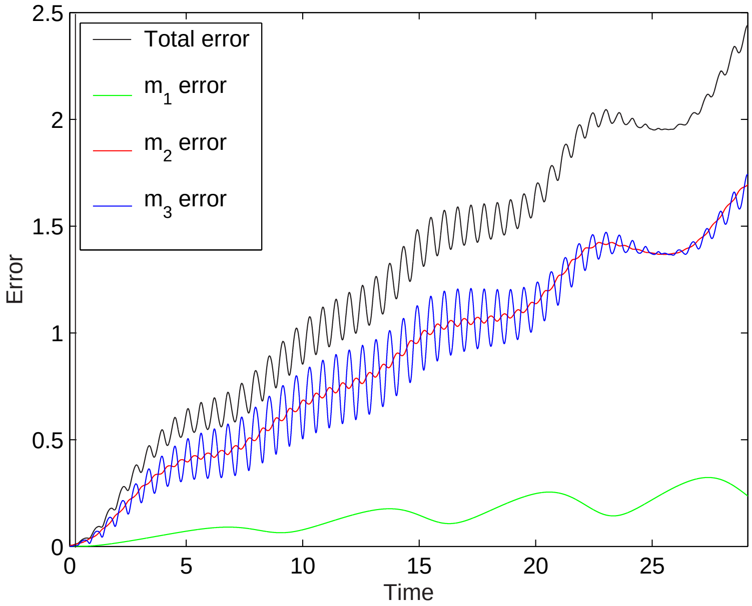


Figure 5.22: Error in the forecast - no observations of m_2; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

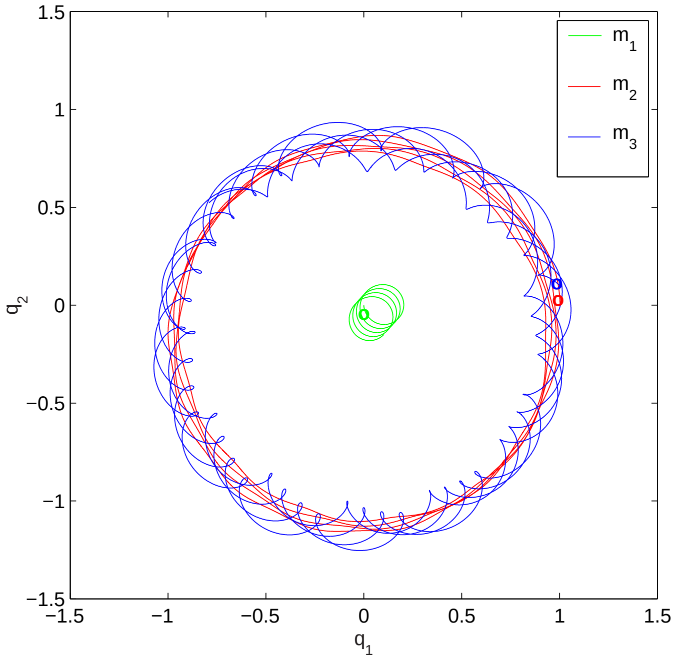


Figure 5.23: Trajectory of the forecast - no observations of m_2; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

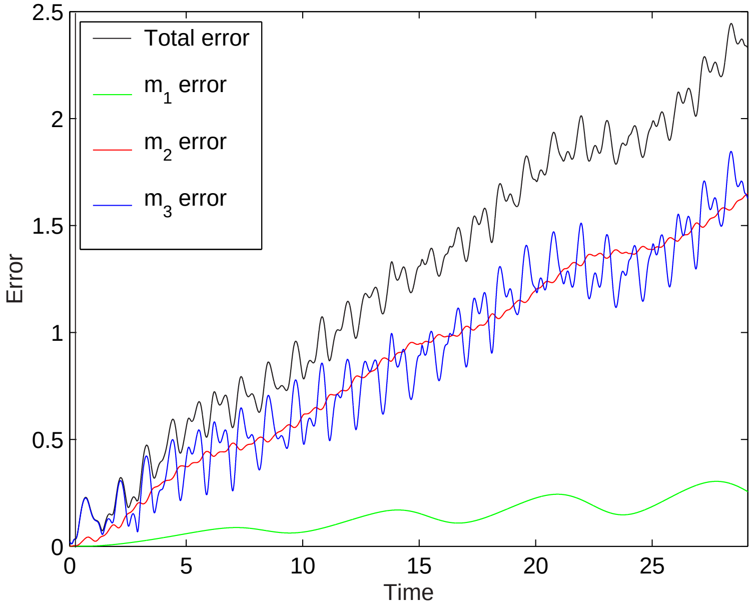


Figure 5.24: Error in the forecast - no observations of m_3; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

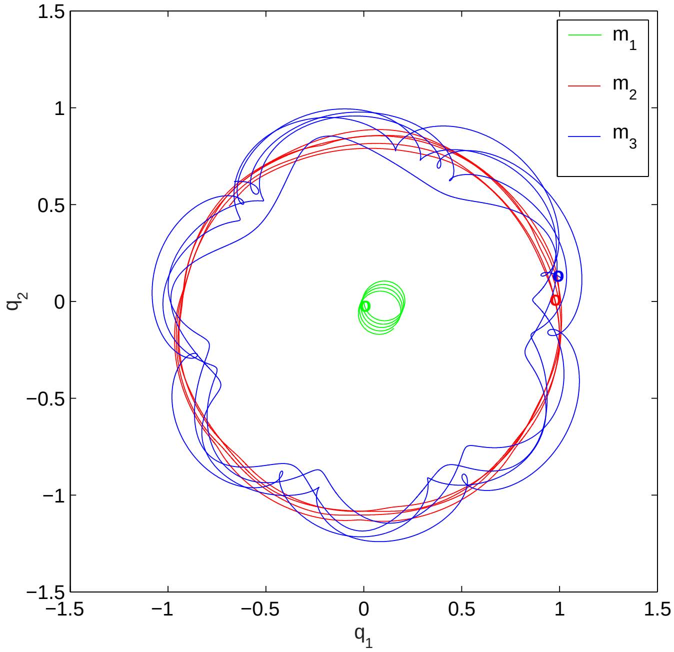


Figure 5.25: Trajectory of the forecast - no observations of m_3; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

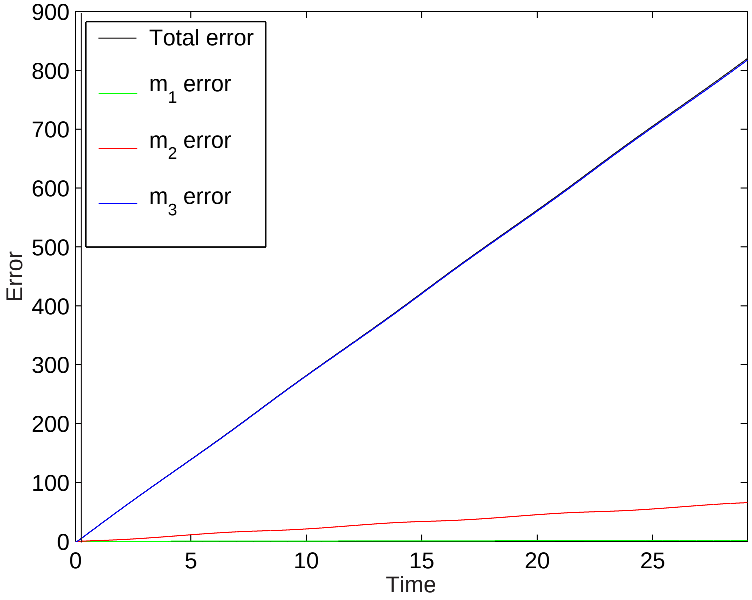


Figure 5.26: Error in the forecast -observations of m_1 only; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

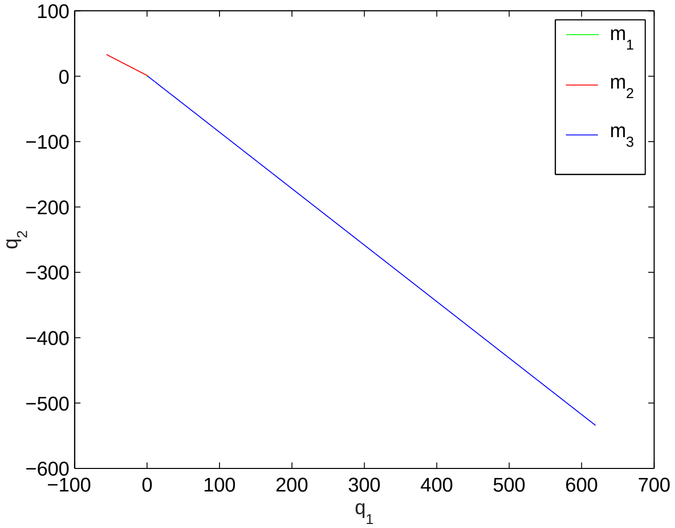


Figure 5.27: Trajectory of the forecast - observations of m_1 only; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

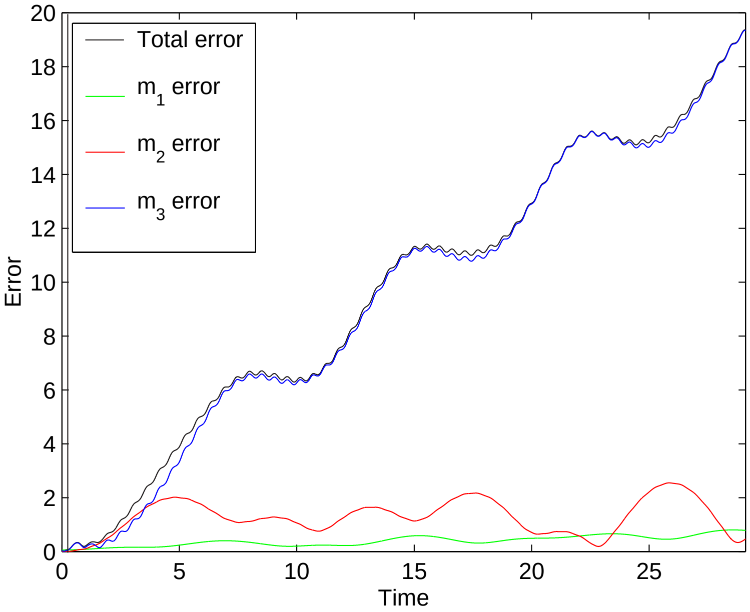


Figure 5.28: Error in the forecast - observations of m_2 only; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

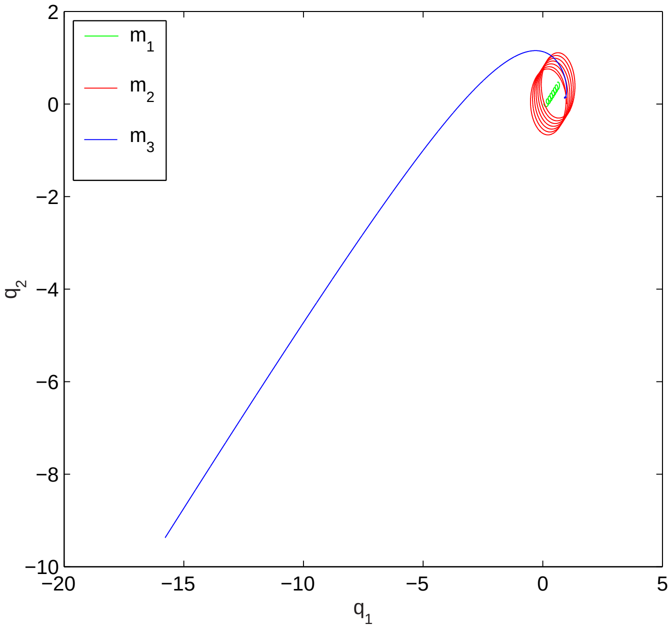


Figure 5.29: Trajectory of the forecast - observations of m_2 only; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

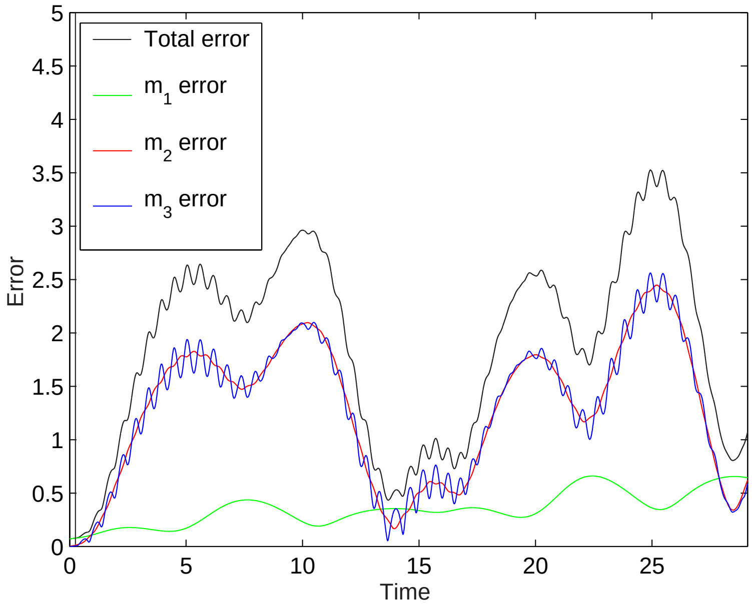


Figure 5.30: Error in the forecast - observations of m_3 only; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

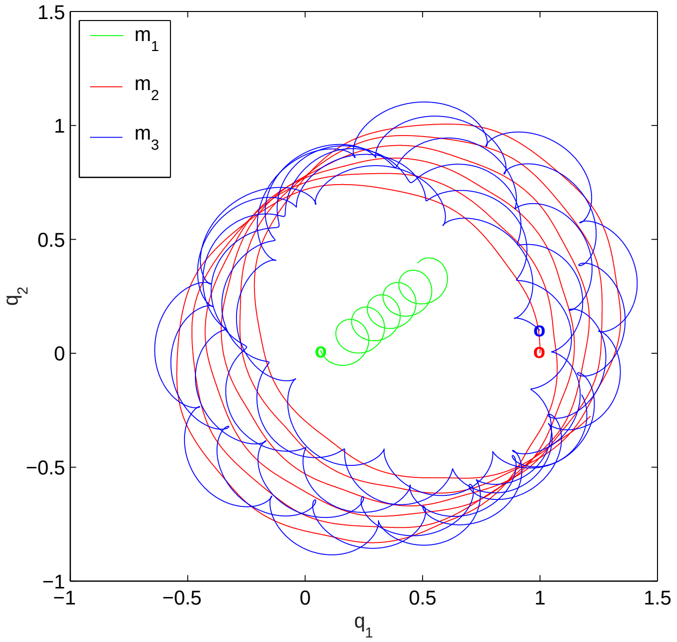


Figure 5.31: Trajectory of the forecast - observations of m_3 only; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

m_2, the characteristic motion of the true solution is not reflected by the analysis. In Figures 5.26 and 5.27 we see that, where we have observations of m_1 only, the analysis has no interaction between the three bodies other than at the initial point. There is not enough information about either the slow or fast motion to reconstruct the solution. In Figures 5.28 and 5.29 we see that the analysis has m_2 in orbit around m_1 but m_3 is not part of the system. Observations of m_2 are thus sufficient to retain the large scale motion. In the final example, shown in Figures 5.30 and 5.31, all three bodies are interacting and the characteristic motion of the true solution is retained. We can therefore conclude that we need observations that give information relating to both scales of motion in order to reconstruct the solution.

5.6 Addition of Weak Constraints

In Chapter 4 we discussed in detail several different additional constraints for the two-body problem. Here we consider the inclusion of similar weak constraints for the three-body problem. We again consider both a perfect background field and a more realistic noisy background field.

5.6.1 Perfect Background

For our first experiments we consider the case where the background is the truth at the initial time. We test these on the example illustrated by Figure 5.19 with a full set of observations at every timestep across the data assimilation window. The observations are created by adding noise to the truth. This has initial condions given by (5.60) to (5.62), (5.66), (5.64) and (5.65).

Our first constraint has the form of a typical background constraint but again, since there are no errors associated with the background the background error covariance matrix is the identity matrix. The constraint has the form

$$J_B = \beta_1(\boldsymbol{x}_b(t_0) - \boldsymbol{x}(t_0))^T(\boldsymbol{x}_b(t_0) - \boldsymbol{x}(t_0)), \tag{5.79}$$

where $\boldsymbol{x}_b = (\boldsymbol{Q}_b^T, \boldsymbol{P}_b^T)^T$ and β_1 is the weighting of the term. As for the two-body problem this constrains the analysis to be close to the background state. Figure 5.32 illustrates the results for several different values of β_1. We can see that for smaller values of β_1 the large scale oscillations are reduced but the small scale oscillations are not. As β_1 increases the amplitude of both the large and small scale oscillations are reduced but the error in the overall solution is still increasing.

The second constraint is the energy constraint, which constrains the energy of the analysis to be close to the energy of the background. This has the form

$$J_E = \beta_2\left(E(\boldsymbol{x}_b(t_0)) - E(\boldsymbol{x}(t_0))\right)^2, \tag{5.80}$$

where $E(\boldsymbol{x})$ is the total energy defined by (5.7). The effect of J_E is demonstrated in Figure 5.33. From this figure we can conclude that the energy constraint does not reduce the amplitude of either the small or large scale oscillations but for larger values of β_2 the overall error is no longer increasing.

We can see that in all cases the constraints improve the forecast. However the behaviour of the solutions is somewhat different. In Figure 5.32 we see that the error in the forecast is reduced but, for all β_1, the error is still increasing for the length of the forecast. We can also see that both the large scale and small scale oscillations are considerably smaller than the unconstrained case. By contrast in Figure 4.17 there is also a reduction in the error, however in this case for large values of β_2 the error begins to decrease. We also observe that amplitude of the small scale oscillations is not reduced.

In order to better understand these results we need to try and consider the two scales of motion separately. This configuration of the three-body problem can be approximated as two two-body problems, that of the centre of mass of m_2 and m_3 orbiting m_1, and m_2 and m_3 in orbit with each other. If we could separate the error trajectory into these two components then we can use the results from Chapter 4 to help interpret the results we have in this chapter.

5.6.2 Separation of Timescales

As we have discussed, one feature of the three-body system we have considered is the distinct timescales within the motion. We have the large scale, slow motion of the two bodies, m_2 and m_3, in orbit around m_1 and the fast motion of the two smaller bodies in mutual orbit. It is difficult to completely separate these two scales of motion. We attempt to do this by considering the evolution of the centre of mass of m_1 and m_2 as they orbit m_1.

The position, $\boldsymbol{q}_{CM}$, and momentum, $\boldsymbol{p}_{CM}$ of the centre of mass are given by

$$\begin{aligned} \boldsymbol{q}_{CM} &= \frac{m_2\overline{\boldsymbol{q}}_2 + m_3\overline{\boldsymbol{p}}_3}{m_2 + m_3} && (5.81) \\ \boldsymbol{p}_{CM} &= \overline{\boldsymbol{p}}_2 + \overline{\boldsymbol{p}}_3, && (5.82) \end{aligned}$$

where $\overline{\boldsymbol{q}}_2 = (q_3, q_4)$, $\overline{\boldsymbol{q}}_3 = (q_5, q_6)$, $\overline{\boldsymbol{p}}_2 = (p_3, p_4)$ and $\overline{\boldsymbol{p}}_3 = (p_5, p_6)$.

We calculate the evolution of the position and momentum of the centre of mass, $\boldsymbol{Q}_{CM}$ and $\boldsymbol{P}_{CM}$, given by the discrete model. We split the total error of the original system into its fast and slow components by defining the slow scale component in the following way,

$$Er_{SLOW} = ||\Delta \boldsymbol{X}_{SLOW}|| \tag{5.83}$$

where

$$\Delta \boldsymbol{X}_{SLOW} = \begin{pmatrix} 2(Q^{tr}_{CM_1} - Q_{CM_1}) \\ 2(Q^{tr}_{CM_2} - Q_{CM_2}) \\ P^{trr}_{CM_1} - P_{CM_1} \\ P^{t}_{CM_2} - P_{CM_2} \\ Q^{tr}_1 - Q_1 \\ Q^{tr}_2 - Q_2 \\ P^{tr}_1 - P_1 \\ P^{tr}_2 - P_2 \end{pmatrix},$$

and $\boldsymbol{Q}^{tr}_{CM}$, $\boldsymbol{P}^{tr}_{CM}$ give the evolution of the centre of mass of the true solution. In this measure of the error we include the error on m_1 in addition to the error on the centre of mass. We multiply the centre of mass position components of $\Delta \boldsymbol{X}_{SLOW}$ by two. This is to account for the fact that when we calculate the total error, we have two large scale motions, m_2 and m_3, whereas the centre of mass has only one. The momentum components are not multiplied by two since the momentum of the centre of mass is given by the sum of the momentum of the two bodies. The small scale motion is simply the difference between our original total error and Er_{SLOW}. We can see in Figure 5.34 that this splits the original error trajectory into the different scales.

We now use these measures to try and gain a greater understanding of the effect of our constraints. We begin by considering the slow component of the error for each of our two constraints. In Figure 5.35 we show the effect of the background term. We see that as we increase β_1 the error in the slow component decreases. For very large β_1 we see there is virtually no error in the slow component. By contrast, in Figure 5.36, where we illustrate the effect of the energy constraint, we see that while there is a significant reduction in the slow component of the error, even for large values of β_2 this is still growing. Additionally by comparing the two constraints, we can see that while the magnitude of the large scale oscillations is reduced where we have imposed the background constraint, it has actually *increased* where we have applied the energy constraint.

We now consider the small scale, fast component of the error. We consider only the effect of the largest value of β. In Figure 5.37 we can see that where we have $\beta_1 = 10^9$, the fast component of the error is increasing with time but the amplitude of the observations has been significantly reduced. In Figure 5.38 we show the corresponding figure for the energy constraint. Here we see that size of the oscillations has reduced somewhat but there is a periodic nature to this, with the period being equivalent ot the length of one large orbit. We also see that there is initially an overall growth of this error component but this begins to decrease towards the end of the forecast.

We can understand what effects these constraints have on the structure of the data assimilation solution by considering the analogous results in the two-body problem. We recall that where we observed a phase error this was a result of comparing two trajectories with a different period, where we observed oscillations in the error this was due to comparing different eccentricities.

We begin by considering the background constraint. We first consider the effect on the slow component. Here we see that the background constraint has a very significant effect on this part of the solution. We observe that there is no growth in the error, and the oscillations are no longer evident. From this we can summise that this constraint is having an effect on the phase error and thus the data assimilation solution has the same large scale period as the truth. We can also infer that the eccentricities of the two are the same. We now consider the effect on the fast motion. We can see that while the magnitude of the oscillations has decreased, there is a phase error between the data assimilation solution and the truth. Thus we can say that the background constraint is not imposing a constraint on the period of the small scale orbit, but is imposing a constraint on its eccentricity.

We now consider the effect of the energy constraint, initially on the large scale orbit. We do not repeat the results that we observed with the two-body problem. Here we see that there is still a phase error, thus we are not constraining the period of the data assimilation solution. With the small scale component we see that the size of the oscillations is somewhat reduced and the overall error begins to decrease. We also observe that the small scale oscillations exhibit a periodic behaviour.

These experiments illustrate that additional constraints can change the analysis in different ways. However, in reality the background would not be the truth and would have some associated error. In the next section we impose constraints on a noisy background field.

5.6.3 Noisy Background

In this section we add the same form of constraints to the cost function as in the previous section, however we now add noise to the background. We now use an example with sparser observations than in the previous example. The observations set is the same as for the case illustrated by Figure 5.19 however we only assimilate observations every ten timesteps. We add weak constraints as in the previous section, however the background state is no longer perfect. This is achieved by adding random noise with a Gaussian distribution to the truth at the initial time. Due to the fact that the magnitude of the variables is somewhat different we use different values of the variance for each one as we did to create the noisy observations. Again we have $\sigma^2_{Q_1,Q_2,P_1,P_2,P_3,P_4} = 10^{-4}$, $\sigma^2_{Q_3,Q_4,Q_5,Q_6} = 10^{-3}$ and $\sigma^2_{P_5,P_6} = 10^{-5}$.

The background constraint thus becomes,

$$J_B = (\boldsymbol{x}_b(t_0) - \boldsymbol{x}(t_0))^T \boldsymbol{B}^{-1}(\boldsymbol{x}_b(t_0) - \boldsymbol{x}(t_0)), \tag{5.84}$$

where

$$\boldsymbol{B} = \begin{pmatrix}
\sigma^2_{Q_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \sigma^2_{Q_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \sigma^2_{Q_3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \sigma^2_{Q_4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \sigma^2_{Q_5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \sigma^2_{Q_6} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{P_1} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{P_2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{P_3} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{P_4} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{P_5} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma^2_{P_6}
\end{pmatrix}$$

The energy constraint now has the form

$$J_E = \frac{1}{\sigma^2_E}\left(E(\boldsymbol{x}_b(t_0)) - E(\boldsymbol{x}(t_0))\right)^2, \tag{5.85}$$

where the variance σ^2_E is defined by

$$\sigma^2_E = \nabla_{\boldsymbol{x}_b} E(\boldsymbol{x}_b(t_0)) \cdot \boldsymbol{\sigma}^2_{\boldsymbol{x}_b} \tag{5.86}$$

and $\boldsymbol{\sigma}_{\boldsymbol{x}_b}^2 = (\sigma_{Q_1}^2, \ldots, \sigma_{Q_6}^2, \sigma_{P_1}^2, \ldots, \sigma_{P_6}^2)^T)$.

The effect of these two different constraints are illustrated by Figure 5.39. This illustrates that both constraints imposed alone have a positive effect on the results, but the addition of both constraints together significantly reduces the error in the forecast. This is due to the fact that we are providing the assimilation algorithm with additional information. The resulting analysis must lie close to the background field but can only do this where it also conserves the energy.

5.6.4 Separation of Timescales

We again consider the separation of the two timescales for the error trajectories given in the previous examples where a noisy background was used. We use the same method of splitting the error into the fast and slow components as demonstrated in Section 5.6.2.

In Figure 5.40 we illustrate the effect on the slow scale of the three different configurations; the background constraint only, the energy constraint only and both imposed together. The overall results are similar to those illustrated by the total error. We find that the energy constraint gives a slightly better result than the background, but imposing both together significantly reduces the error.

In Figure 5.41 we show separately the effect on the small scale for the three different cases. We can see clearly that the background constraint produces small oscillations that are constantly growing in size. By contrast, the energy constraint again produces periodic behaviour in these oscillations. We can also see that for the third case, where we have observed a significant reduction in the overall error, by the end of the window the oscillations are larger than in the other two cases.

5.6.5 Example: Incomplete observations

In this section we apply our two different constraints to the cases in Section 5.5.1 where the characteristic behaviour of the true solution was not exhibited by the analysis. We consider whether our constraints can provide sufficient additional information to construct an analysis that has the same properties as the truth. We begin by considering the case where we assimilated only observations of the largest body, m_1. We now add our weak constraints to the cost function for this example. Therefore the effect of these constraints, shown in Figures 5.42 to 5.44, can be observed by comparison with the unconstrained case illustrated by Figure 5.26. In Figure 5.42 we illustrate the effect of including J_B. Here we can see that the error is considerably smaller than for the unconstrained case. The overall behaviour demonstrated by the forecast has not changed significantly, however the three bodies interact for longer before moving away from each other.

We now consider the addition of the energy constraint, J_E, to the cost function. The effect of this is illustrated by Figure 5.43. The constraint has a much more significant impact than J_B. The error is an order of magnitude smaller than in the previous example. We also observe that there is now interaction between m_1 and m_2 for the duration of the forecast. However the third body is ejected from the system and we do not observe the stable behaviour inherent in the true solution.

We finally consider the addition of both our constraints to the system. In Figure 5.44 we see that the error is almost ten times smaller than in the previous case, where we included the energy constraint only. We also observe, by considering the trajectories shown in Figure 5.44(b), that the three bodies interact with each other for the duration of the forecast. They do not, however, reproduce similar behaviour to the true solution, where m_3 orbits m_2 and they both orbit m_1.

We now consider the case where only observations of m_2 were assimilated. Thus we are comparing Figures 5.45 to 5.47 with the unconstrained example illustrated by Figure 5.28. In Figure 5.45 we show the effect of the background constraint on this example. We see that this constraint has provided enough information such that the analysis exhibits behaviour characteristic of the truth.

We now consider the addition of the energy constraint only. We can see in Figure 5.46 that this constraint is not as effective as the background constraint. The error is almost four times as large as in the previous case. We can also see that while the three bodies are interacting for the full length of the forecast, this interaction does not reflect the truth. Here we observe that while m_2 and m_3 are both in orbit around m_1, m_3 is not orbiting m_2.

In the final example we include both the energy and the background constraints. As we might expect, Figure 5.47 shows that we again produce the behaviour that is observed in the true solution. The addition of the energy constraint, however, acts to reduce the overall error.

We saw in Section 5.6.3 that the addition of our weak constraint had the effect of reducing the error in the solution. In this section we note that, where observations prove insufficient, the constraints can have a significant and positive effect. We cannot say which of the constraints is more effective. In the first example the energy constraint had a greater effect on the error, however, where we had observations of m_2 only, the background constraint was able to reconstruct the correct behaviour. We can conclude, however, that the inclusion of *both* constraints significantly improves the analysis and forecast where the observations alone have failed to produce the desired behaviour.

5.7 Summary and Conclusions

In this chapter we have discussed the three-body system, a more general problem than that considered in Chapter 4. As in this previous chapter we derived the continuous canonical equations and stated the conservation properties of the system. We then derived the discrete model equations using the Störmer-Verlet method and tested the resulting model. We showed that it preserved the inherent characteristics of the original system, specifically the conservation of the Hamiltonian - the total energy of the system.

We also considered the implication of our choice of initial conditions. We showed that for this system the behaviour of the three bodies was not always predictable. However we were able to find a configuration that provided an analogy with the atmosphere by exhibiting a separation of timescales. We demonstrated that for this choice of initial conditions the system would remain in its stable orbits but that perturbations to these initial conditions could alter the characteristic behaviour.

In Section 5.4 we discussed the construction and subsequent testing of the the data assimilation algorithm for this problem. We then went on to describe the numerical experiments that were carried out with the completed 4d Var scheme. We initially considered the effect of having an incomplete set of observations. We showed that if observations from only one of the bodies were assimilated, the 4d Var scheme could not reconstruct the two-timescale behaviour of the true solution. Thus the data assimilation scheme, which included only the observation term in the cost function, could not provide enough information. We therefore considered the addition of weak constraints that provide information from a previous forecast. We thus impose them at the initial time only.

The first of these constraints is the background term typical of most 4d Var schemes. For all the examples considered here we showed that this improved the forecast in all cases. The second constraint made use of the conservation property of this system. As we have discussed the three-body system is energy conserving and our numerical model reflects this. We can say, therefore, that the energy conservation is imposed as a strong constraint across the data assimilation window. However we wished to impose this as a further constraint such that the analysis would be weakly constrained to lie close to the energy of the background. In practical terms this implies that the constraint would impose the energy conservation property from one assimilation window to the next.

For the two-body problem we were able to fully understand the implications of the different constraints on the geometry of the resulting solutions. For the three-body problem, due to its more complex nature, this was more difficult. However, we were able to separate the error into a fast and slow component. We found that, in terms of the large, slow motion the background constraint had the effect of constraining both the period and the eccentricity. The energy constraint, however, did not repeat the behaviour observed in the two-body problem. The term did not constrain the period. We saw in the two-body chapter that the period of an orbit was related to its energy and thus an energy constraint affects the period. This behaviour was not repeated here. In this three-body problem we effectively have two two-body problems, but we are only imposing a constraint on the total energy. The system therefore cannot infer how much of the total energy is associated with the large orbit or the small orbit. Thus we find that the large orbit has a different energy and therefore different period to the truth.

When considering the small scale motion, we found that the background constraint did not impose a constraint on the period of the small orbit, although the magnitude of the oscillations was significantly reduced, implying that there was some constraint on the eccentricity. In all cases we found that there was a significant reduction in the error

This proved to be true even with a noisy background field. We observed that the addition of a background term improved the analysis and forecast as expected, since this is the typical formulation of 4d Var assimilation. The addition of the energy constraint alone also improved our results. However the largest reduction in the overall error was found when both constraints were used. As we discussed, the energy contraint does not prevent phase error as it did for the two-body problem as the energy is associated with two different orbits. However in the case of the noisy observation we saw that the addtion of both constraints had a beneficial effect. In this case we are getting information in two different ways, the background constraint may be providing additional information such that the energy constraint is able to provide a more accurate balance between the energies of the two orbits, thus reducing the phase error of both.

We can explain this by considering the minimisation algorithm. As we discussed in Chapter 4, the cost function gradient provides the descent direction for the minimisation algorithm and this must include the gradient of the additional terms. Therefore, the gradient of the cost function will only allow an analysis that not only lies close to the background field but also weakly conserves the energy of the system.

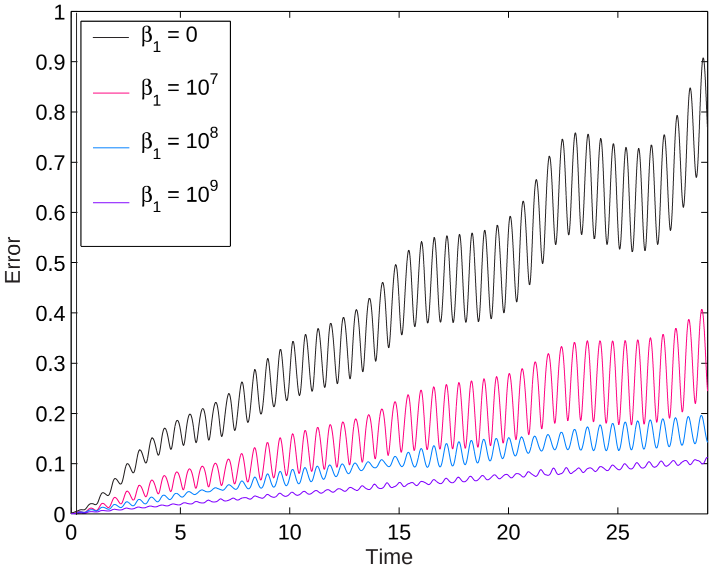


Figure 5.32: Error in the forecast with a perfect background constraint; all observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65), $\beta_1 = 0, 10^7, 10^8, 10^9$

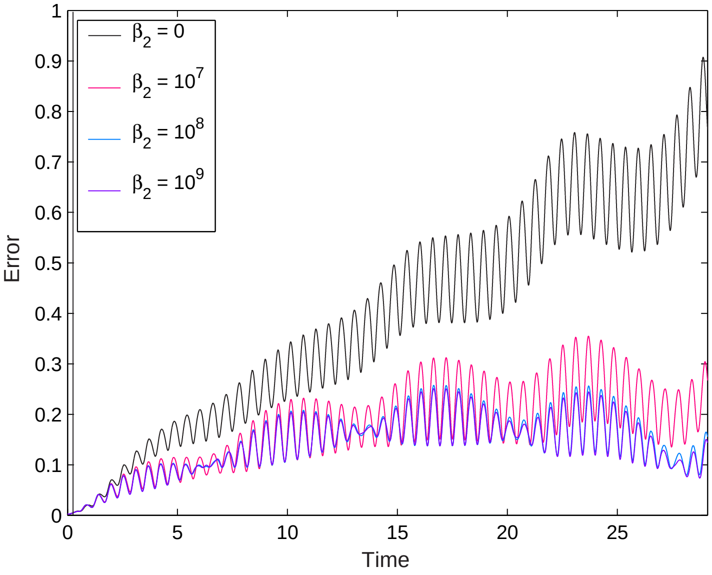


Figure 5.33: Error in the forecast with a perfect energy constraint; all observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65), $\beta_2 = 0, 10^7, 10^8, 10^9$

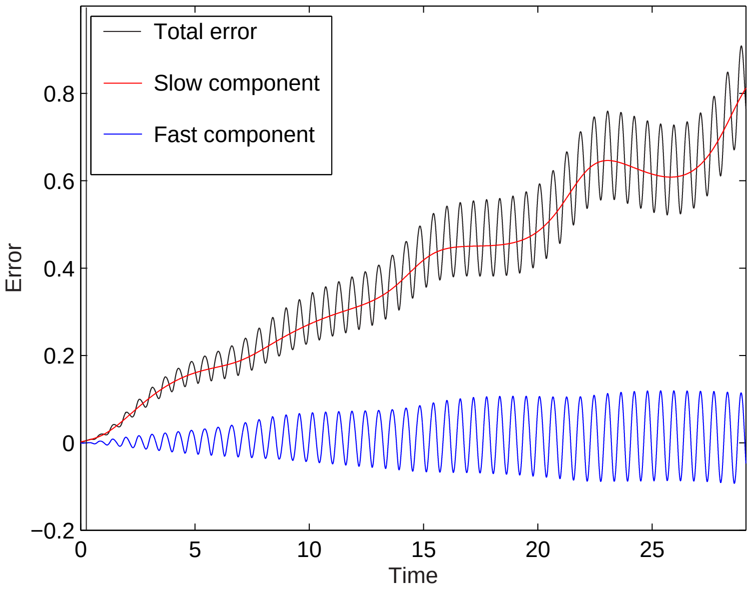


Figure 5.34: Illustration of scale separation of the error trajectory, no constraints, observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

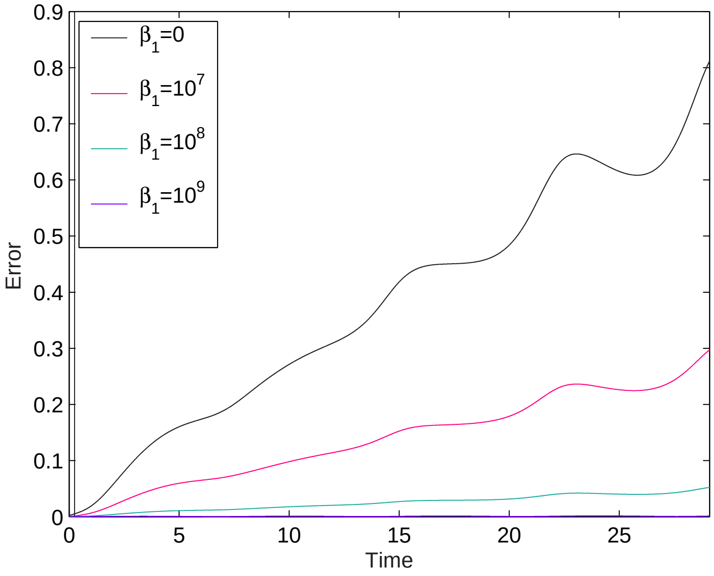


Figure 5.35: Slow scale component of the error in the forecast using a perfect background constraint; all observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65), $\beta_1 = 0, 10^7, 10^8, 10^9$

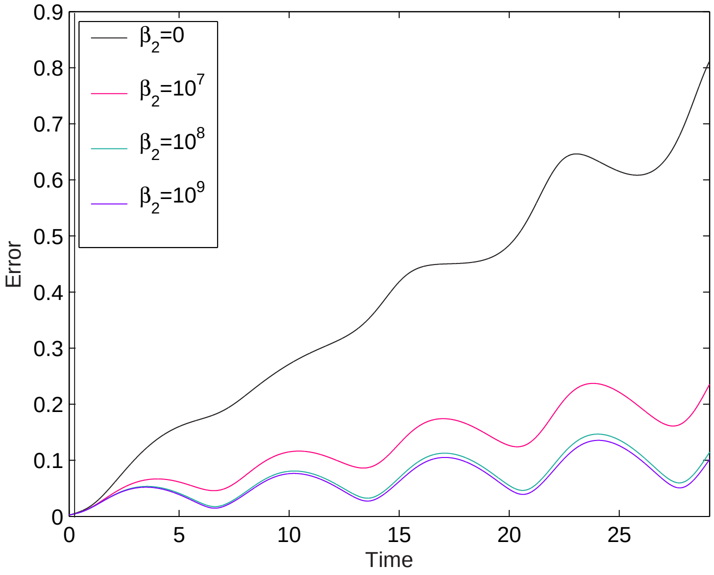


Figure 5.36: Slow scale component of the error in the forecast using a perfect energy constraint; all observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65), $\beta_2 = 0, 10^7, 10^8, 10^9$

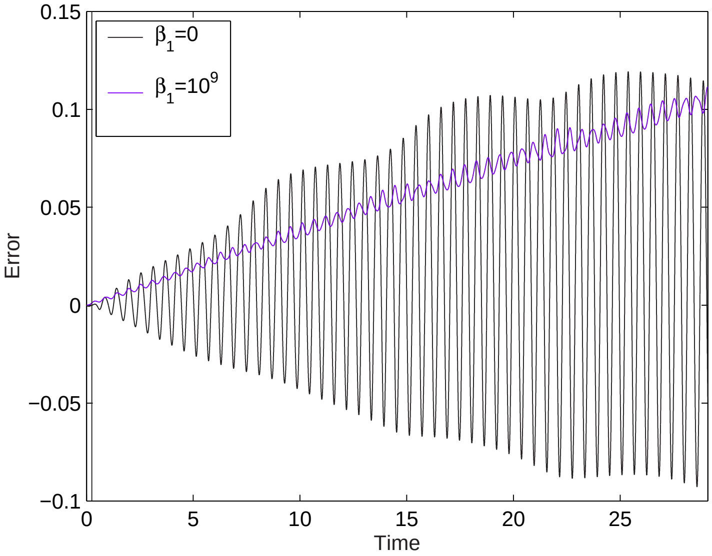


Figure 5.37: Fast component of the error in the forecast using a perfect background constraint; all observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65), $\beta_1 = 0, 10^9$

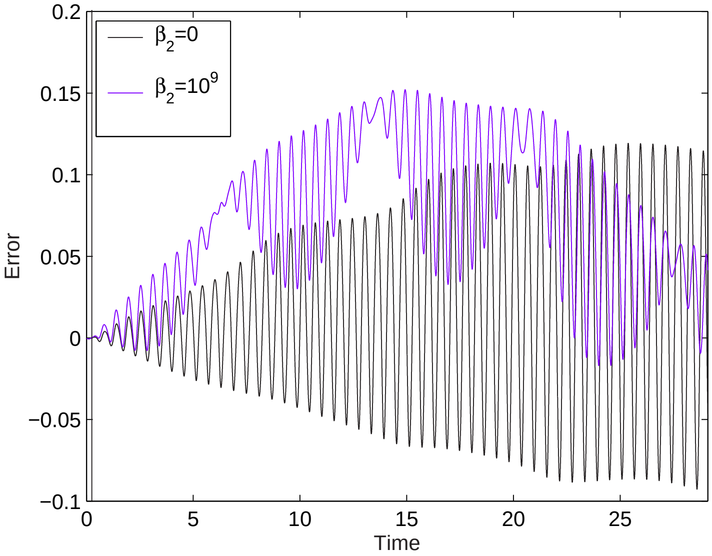


Figure 5.38: Fast component of the error in the forecast using a perfect energy constraint; all observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65), $\beta_2 = 0, 10^9$

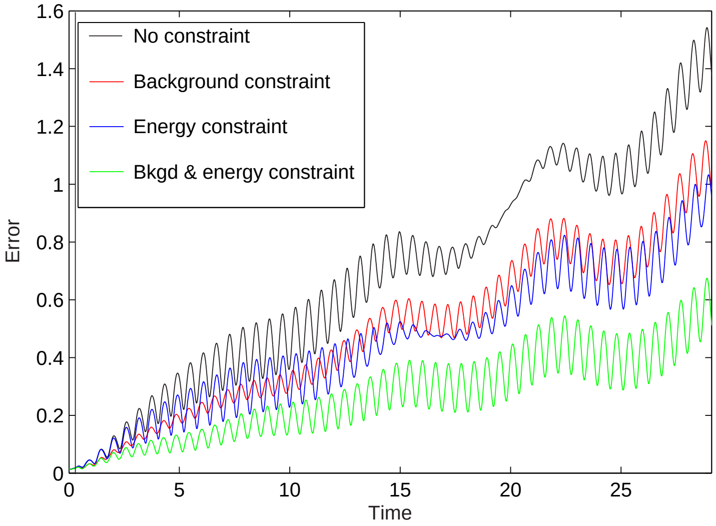


Figure 5.39: Error in the forecast produced by the three-body 4d Var scheme with noisy constraints; observations at every 10 timesteps, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

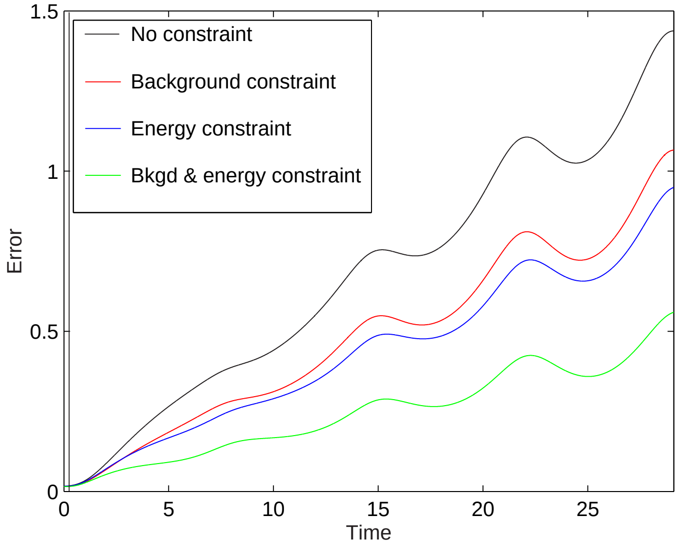


Figure 5.40: Slow component of the error in the forecast using realistic constraints; all observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

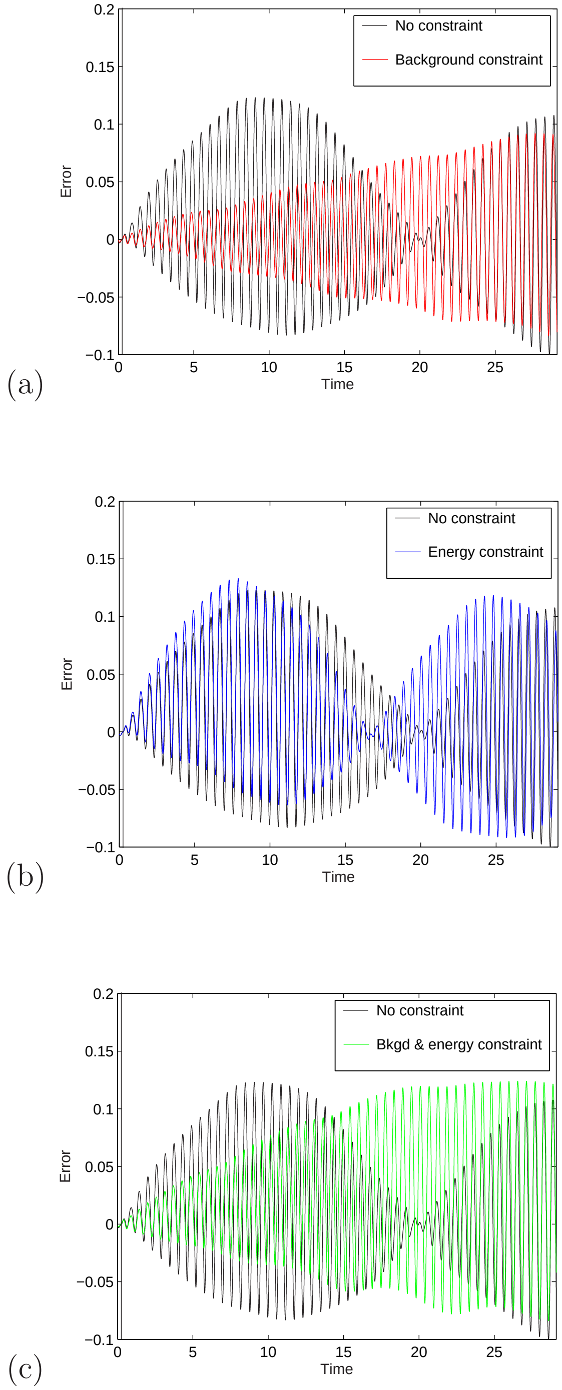
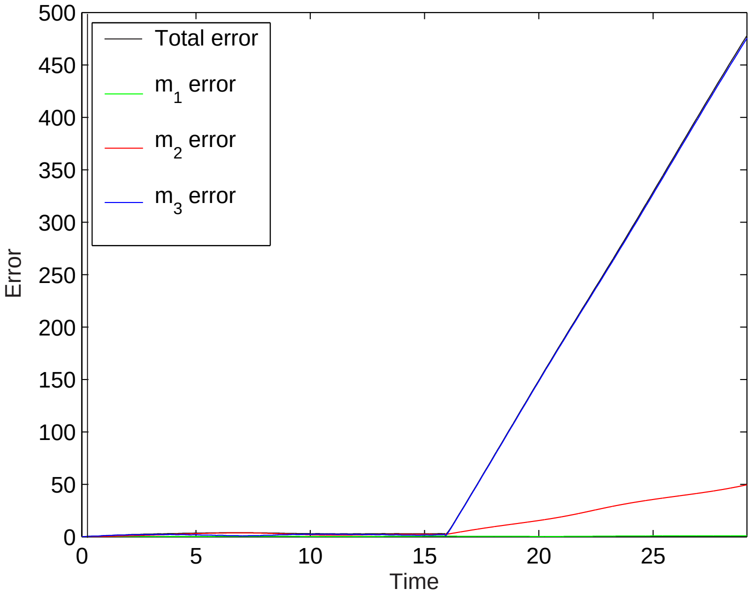
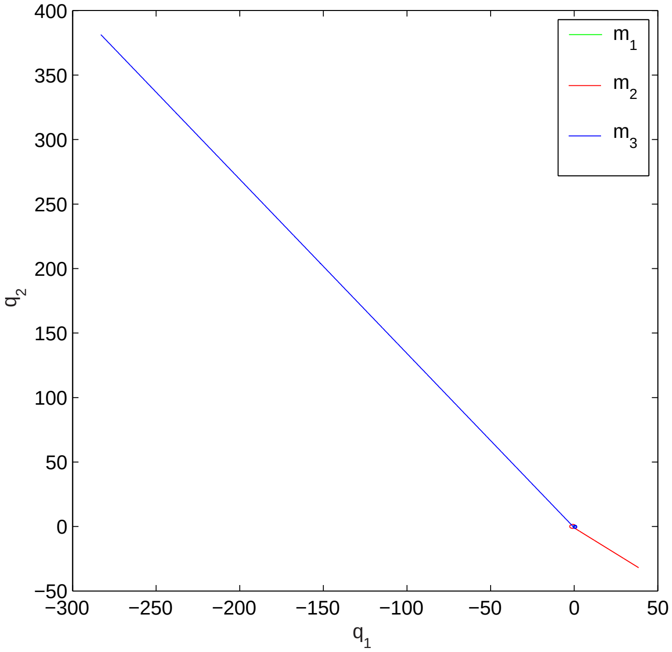


Figure 5.41: Fast component of the error in the forecast using (a) a realistic background constraint, (b) a realistic energy constraint and (c) both constraints; all observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

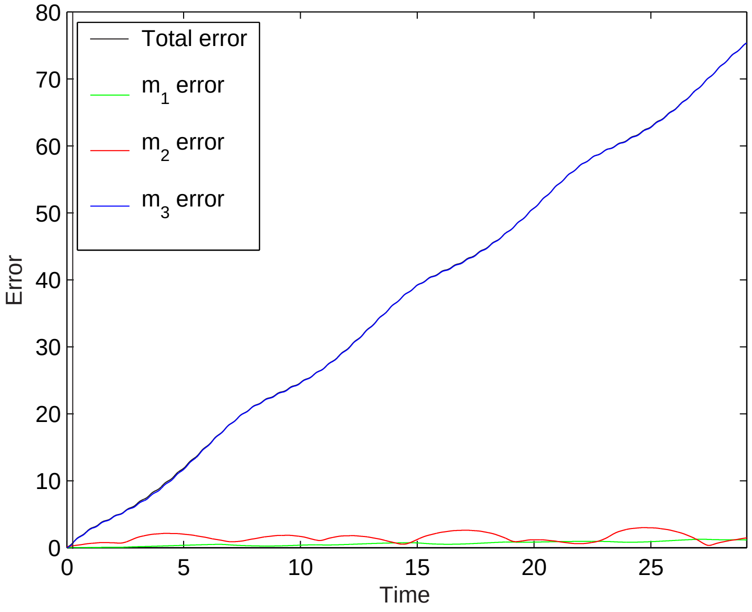


(a)

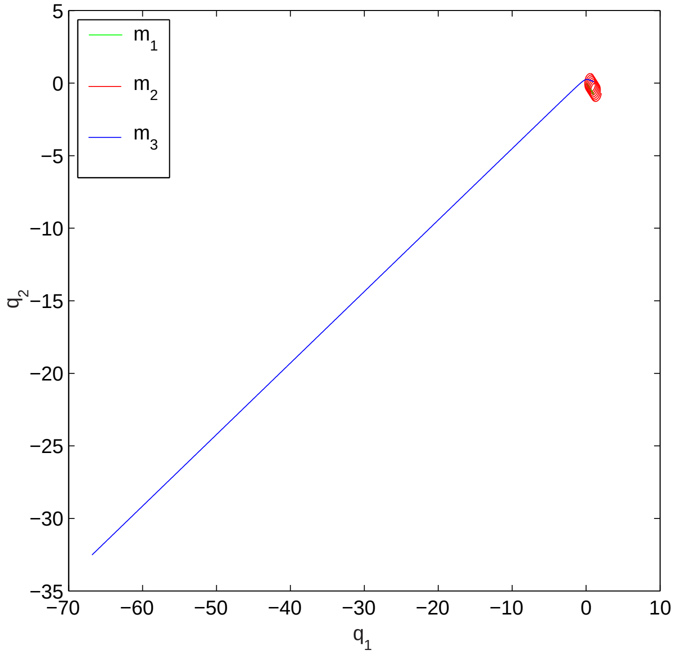


(b)

Figure 5.42: (a) Error and (b) Trajectory of the forecast with observations of m_1 only, background constraint included; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)



(a)



(b)

Figure 5.43: (a) Error and (b) Trajectory of the forecast with observations of m_1 only, energy constraint included; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

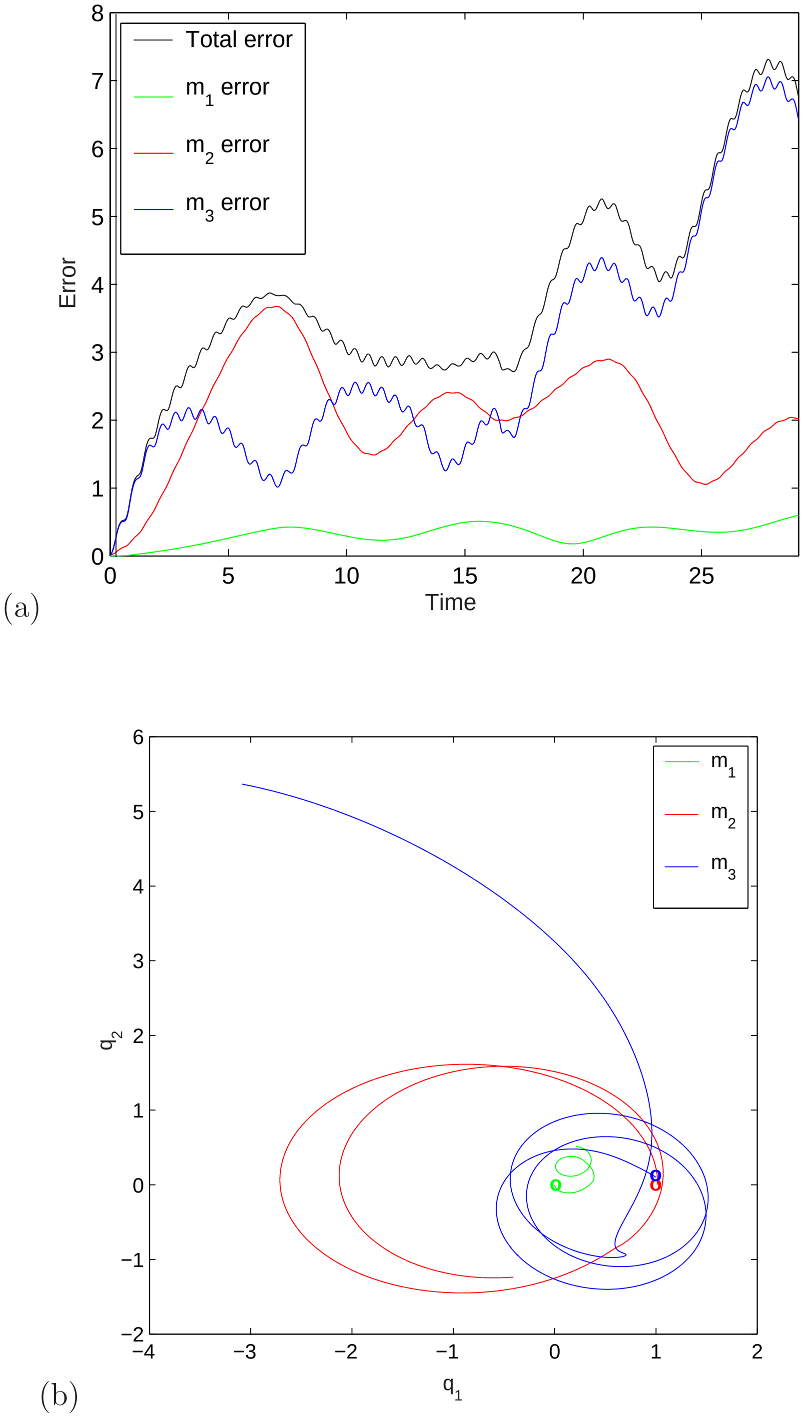
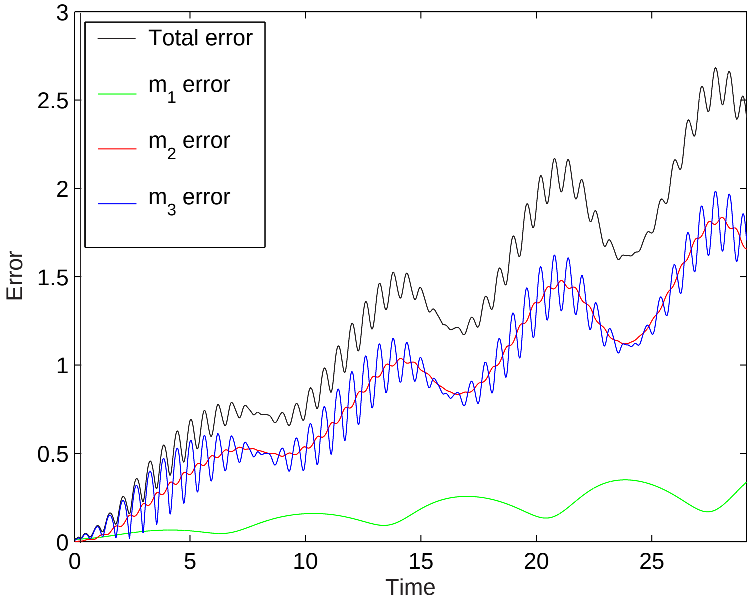
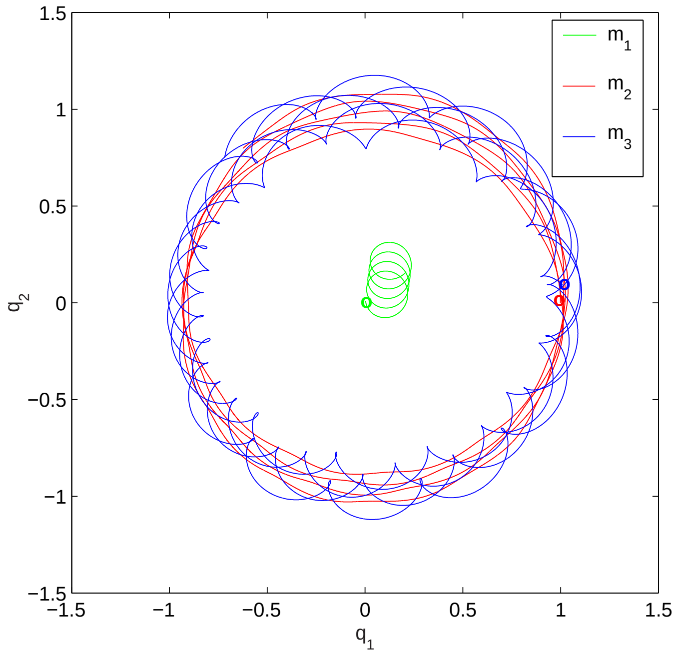


Figure 5.44: (a) Error and (b) Trajectory of the forecast with observations of m_1 only, background and energy constraints included; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)



(a)



(b)

Figure 5.45: (a) Error and (b) Trajectory of the forecast with observations of m_2 only, background constraint included; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

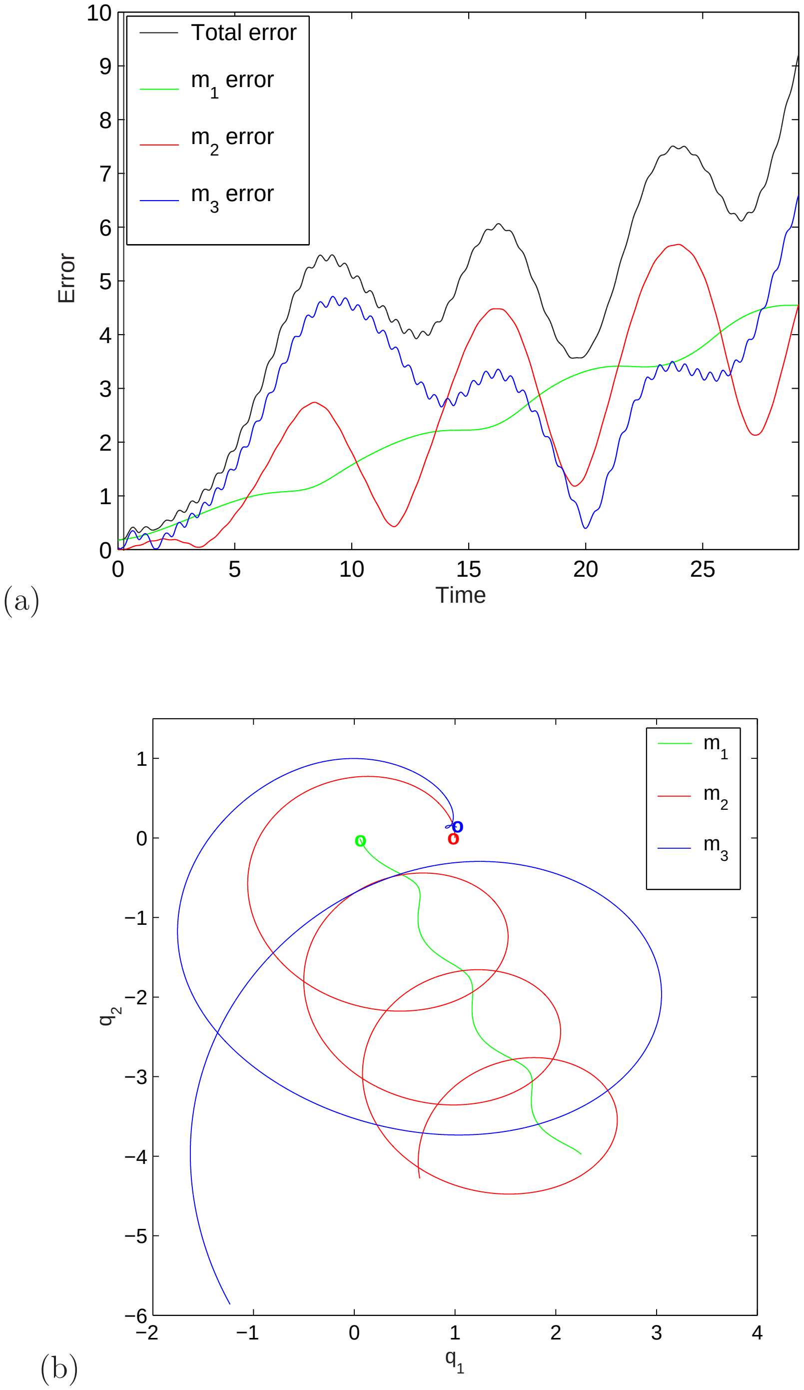
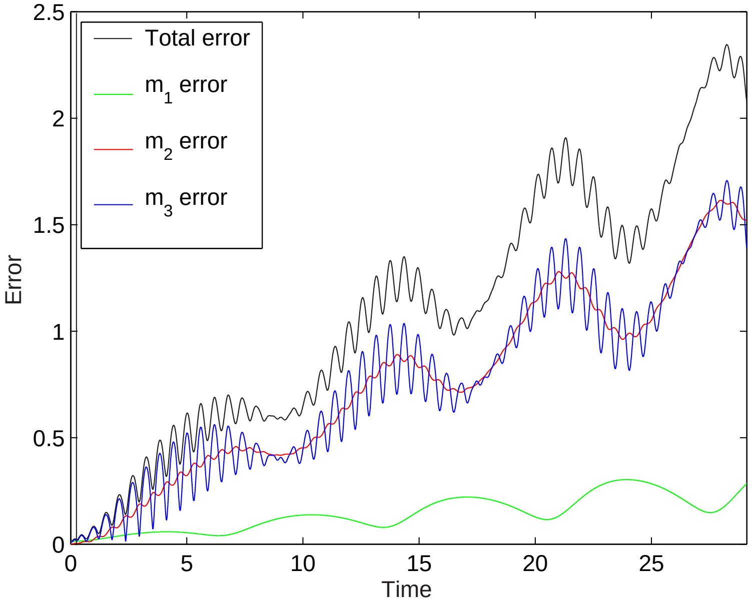
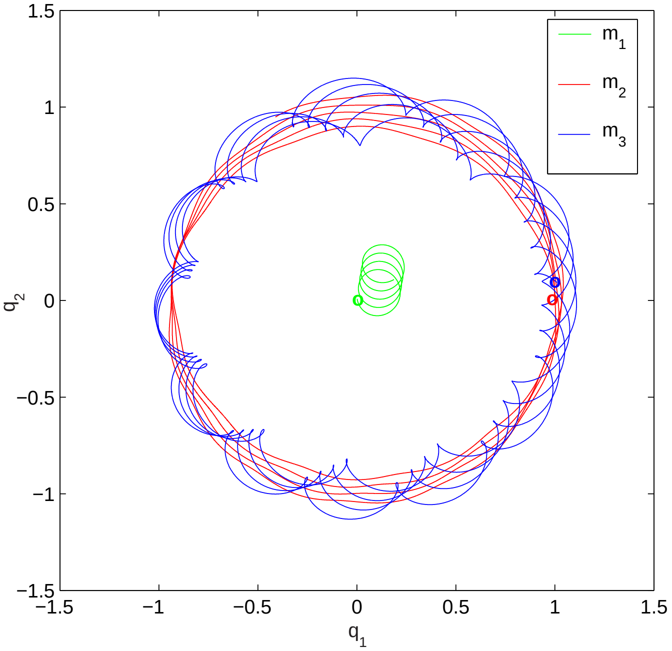


Figure 5.46: (a) Error and (b) Trajectory of the forecast with observations of m_2 only, energy constraint included; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)



(a)



(b)

Figure 5.47: (a) Error and (b) Trajectory of the forecast with observations of m_2 only, background and energy constraints included; observations at every timestep, $T_{DA} = 0.3$, $h = 0.001$, IC: (5.60) to (5.62), (5.66), (5.64) and (5.65)

Chapter 6

Conclusions and Further Work

6.1 Conclusions

This thesis has been concerned with the construction of data assimilation schemes for Hamiltonian systems. We have demonstrated that the addition of weak constraints to the cost function produces an improved analysis and subsequent forecast. Specifically we have illustrated the effect of including a constraint that weakly imposes the conservation properties of the system at the initial time. We have seen that the addition of such constraints to the typical 4d Var system, where $J = J_b + J_o$, results in a significant reduction to the error in the analysis and forecast.

We began in Chapter 2 by introducing Hamiltonian problems and giving details of the mathematics for such systems. We demonstrated the significance of the conservation properties which are intrinsic to Hamiltonian systems. We showed that symplectic techniques are particularly successful for modelling such problems as they prove to be excellent at conserving the inherent properties of the system.

In Chapter 3 we introduced the concept of data assimilation. We discussed both the typical formulation of 4d Var and the addition of weak constraints to the system.

Chapters 2 and 3 provided the information necessary to construct a 4d Var scheme for a Hamiltonian problem. In Chaper 4 we made use of this to construct a data assimilation algorithm for the two-body problem. We introduced the reduced two-body problem that describes the system as one body of reduced mass orbiting a fixed body of total mass. This formulation results in a system that can be qualitatively described using Kepler's Laws. These relate to the symmetry and geometry of the problem and thus enhanced our understanding of our later results.

Having described the problem qualitatively we then derived the continuous canonical equations and the resulting numerical model for the two-body problem. This model was used to construct the 4d Var scheme for the system. We initially considered a cost function containing the observation term only. We began by considering the effect of assimilating different observation sets. We compared the assimilation of a full state vector, observations of position only and observations of momenta only. We then assessed the effect of using sparse observations, assimilating eight observations in each case. We found that where we assimilated fewer observations at more frequent intervals, the analysis and forecast were improved.

From these we selected an example where the error in the forecast was increasing with time to provide a test case for our future experiments. Our aim was to reduce this error by adding various extra terms to the cost function. The first of these was a typical background term. This constrains the state vector of the analysis to lie close to the background state vector. The remaining constraints exploited the conservation properties of the system. The first of these weakly constrained the energy of the analysis to be close to the energy of the background. Thus we effectively, weakly impose energy conservation from window to window. Similarly we considered a constraint on the angular momentum.

When adding our constraints, we looked at two different scenarios. Initially we used a perfect background field. While unrealistic, this set up allowed us to gain an insight into the effect of the different constraints on the geometry of the solution. We found that all constraints provided an improved analysis and forecast but the results were somewhat different. The energy and angular momentum constraints reduced the phase error and thus we could infer that they imposed a constraint on the semi-major axis and period of the orbit. This is due to the fact that any orbits with the same semi-major axes will also have the same energy and angular momentum. We also found that the effect of these constraints was virtually identical. The background constraint, however, reduced the amplitude of the oscillations in the error but not the phase error. From this we could deduce that this term has the effect of constraining the eccentricity of the analysis.

We finally considered the addition of noise to the background field. We found again that all three constraints provided an improvement to the analysis and forecast. Here the typcial formulation of the background constraint produced the best results. However when we added the energy constraint to the observation term *and* the background term this further reduced the error. Similar results were found when considering the angular momentum constraint. We also considered the effect of adding all the additional constraints. While this improved the solution further, the additional angular momentum term had the same effect as if we had doubled the weighting on the energy constraint. This implies that we gain no different information about the background field from the angular momentum constraint.

The next stage of this thesis was to consider a more general problem. In Chapter 5 we again use the information from Chapters 2 and 3 to construct a 4d Var scheme for the three-body problem. As for the two-body problem we derived the continous equations for this energy conserving, Hamiltonian problem and constructed the numerical model. For the three-body problem the choice of initial conditions was important. The nature of the system means that the behaviour of the three bodies is not always predictable. Our final choice of initial conditions produced a Sun-Earth-Moon type system with the large scale motion of the second and third bodies orbiting the first and the small scale of the third body in orbit around the second.

We also tested the stability and sensitivity of this configuration. We showed that while the system can be considered stable for the length of our model run, perturbations to the initial conditions could result in completely different behaviour. Thus our assimilation scheme could produce an analysis that does not resemble the true solution.

For our numerical experiments using the three-body 4d Var scheme, we initially looked at the effect of assimilating different observation sets. We found that if we only assimilated observations from one body, the resulting assimilation solution exhibited the behaviour that was significantly different to the truth.

As we found scenarios where the data assimilation scheme failed, we needed to again consider the addition of further constraints. For the three-body problem we investigated the effects of the background constraint and an energy constraint, both applied at the initial time. We began by finding an example where, without any constraints, the characteristic behaviour of the analysis is similar to the truth but has an error that is increasing with time. This is similar to the case we used in the two-body problem and thus were able to infer certain effects of the constraints by analogy with the two-body problem.

The first experiments used a perfect background field. We found that both constraints improved the analysis and forecast in different ways. The background constraint reduced the size of both the small and large scale oscillations. By comparing this with the two-body results we thus deduced that we were constraining the eccentricities of all the orbits. With the energy constraint we observed that the amplitude of the oscillations was not reduced therefore the orbits of the analysis have different eccentricities to those of the true solution. We noted that the large scale behaviour did not mirror that of a two-body problem and phase error was observed. This is because there are two orbits each with an associated energy associated. Since the constraint is on the total energy, the assimilation scheme cannot determine how much energy each orbit has alone. Thus our solution does not have the same large scale period as the truth.

We next looked at the addition of noise to this background field. Both constraints alone reduced the overall error, imposing both produced an even better analysis. Here the background term provides further information to the system and thus may positively affect the balance of the energy between the two scales, producing orbits that have energies and thus periods that are closer to the truth.

We applied these constraints to the cases discussed previously, where the observation term alone could not reproduce the characteristic behaviour of the true solution. We demonstrated that the additional constraints significantly increased the interaction of the three bodies. We thus deduced that our additional constraints can not only reduce the error in our analyses and forecasts, in some cases they can also provide enough additional information to produce the appropriate behaviour when the observation term alone cannot.

In our work on the two-body and three-body problems we have shown that the addition of a weak constraint conserving the energy can improve our results further than imposing the background term alone. By imposing this constraint we are providing the data assimilation scheme with further information. We are saying that we want the analysis to lie close to the background but only in such a way that the energy is weakly conserved between the background field and the anaysis. The gradient of the cost function includes the gradient of this energy term. This gradient provides the descent direction for the minimisation algorithm. The gradient will therefore only allow an analysis that satisfies both weak constraints.

In this thesis we have found that adding to the cost function a weak constraint that reflects the conservation property of the system, provides an improved data assimilation solution and forecast. We have seen that this term provides additional information about the structure of the solution. This may have an implication for numerical weather prediction (NWP) models, since these exhibit conservation of potential vorticity. Similar constraints could be investigated in NWP data assimilation schemes to see if the results demonstrated in this work can be extended to more complex systems.

6.2 Further Work

The 4d Var data assimilation algorithm for the three-body problem will provide a useful testbed for many research questions regarding data assimilation for Hamiltonian problems. As an extension to this work, we could consider the effect of adding a weak constraint on the angular momentum to the three-body problem.

When we included this term for the two-body problem, it did not provide any further information than the energy constraint. However, the three-body problem has more degrees of freedom, and so the addition of such a constraint may have a greater effect on the analysis and forecast. It might also prove useful to transform this system so that the origin lies at the centre of mass. This would therefore remove from consideration any error that is associated with the transverse motion of the entire system. We would thus be able to see more clearly the effect of the assimilation scheme on the two scales of motion that are of interest.

In addition, as we have discussed throughout this thesis, the three-body problem has some useful features that reflect behaviour in the atmosphere. Over certain timescales the atmosphere exhibits behaviour that is Hamiltonian in nature. In addition we have two distinct scales of motion, the small scale gravity waves and the large scale synoptic systems. Both of these properties can be observed in the three-body problem if a suitable configuration is used.

In this section we will discuss several different areas of research where such an assimilation scheme could be of use.

6.2.1 Fraternal-Twin Experiments

Fraternal-twin experiments use different models to provide the observations. These can often provide a more realistic problem. For example, within data assimilation schemes spurious gravity waves can be spuriously created by the model. These are not features of the real atmosphere which is dominated by the larger scale, smooth, synoptic features. We can consider using models that cannot produce these gravity waves. However, although this would provide a reasonable approximation, it does not allow for the fact that gravity waves are sometimes produced naturally in the atmosphere. We may thus want to use a full model for our data assimilation algorithm to allow for these natural small scale motions. However since the atmosphere is typically smooth we want to assimilate such observations. We can use our three-body model to investigate this. Instead of assimilating observations from the full three-body system we take observations from the corresponding two-body problem that reflects the three-body large scale motion. This would allow us to consider how the model deals with the small scale motion when there are no observations at this scale.

Conversely we could consider the effect of using a two-body data assimilation scheme and taking observations from the three-body system. This would allow us to consider the situation where the model does not allow small scale motion but the observations contain small scale features. Our model may help us understand how these small scale features are interpreted by a data assimilation scheme which cannot produce any small features. We can investigate how this might affect the large scale behaviour in this case.

6.2.2 Imposing Constraints Across the Data Assimilation Window

In this thesis we have considered constraints that are only explicitly imposed at the initial time. Our three-body model would be an appropriate choice when considering the effect of imposing the constraints across the whole of the data assimilation window. In this thesis, we might consider that the energy constraint acts across the entire window since the numerical model is energy conserving, and we are imposing the property at the initial time. However we may consider explicitly imposing this constraint across the assimilation window. This would have the effect of altering the adjoint of the system. The data assimilation scheme, consisting of the forward nonlinear equations and the adjoint equations, is itself Hamiltonian. We could therefore use the model to consider the effect of this energy constraint, an additional Hamiltonian term, might have on the structure of the overall data assimilation scheme.

Bibliography

[1] D. Acheson: *From Calculus to Chaos, An Introduction to Dynamics*, Oxford University Press, 1997

[2] V. I. Arnold: *Mathematical Methods of Classical Mechanics*, Springer-Verlag, 1978

[3] S. Blanes and C. J. Budd: *Explicit, Adaptive, Symplectic (EASY) Integrators Using Scale Invariant Regularisations and Canonical Transformations*, Celestial Mechanics and Dynamical Astronomy, v. 89, 383-405, 2004

[4] O. Bokhove and T. G. Shepherd: *On Hamiltonian Balanced Dynamics and the Slowest Invariant Manifold*, Journal of the Atmospheric Sciences, v. 53, no.2, 276-297

[5] F. Bouttier and P. Courtier: *Data Assimilation Concepts and Methods*, ECMWF Meteorological Training Course Lecture Series, March 1999

[6] C. J. Budd and M. D. Piggott: *Geometric Integration and Its Applications*, from the proceedings of the ECMWF Workshop on Developments in Numerical Methods for Very High Resolution Global Models, 93-117, June 2000

[7] C. J. Budd, B. Leimkuhler and M. D. Piggott: *Scaling Invariance and Adaptivity*, Applied Numerical Mathematics, v. 39, 261-289, 2001

[8] C. J. Budd and M. D. Piggott: *Geometric Integration and Its Applications*, Foundations of Computational Mathematics, Handbook of Numerical Analysis, v. 11, 35-139, 2003

[9] W. C. Chao and L. Chang: *Development of a Four-Dimensional Variational Analysis System Using the Adjoint Method at GLA. Part1: Dynamics*, Monthly Weather Review, 1661-1673, August 1992

[10] P. Courtier, J-N Thépaut and A. Hollingsworth: *A Strategy for Operational Implementation of 4D-Var, Using an Incremental Approach*, Quarterly Journal of the Royal Meteorological Society, v. 120, 1367-1387

[11] M. J. P. Cullen, J. Norbury, R. J. Purser and G. J. Shutts: *Modelling the Quasi-Equilibrium Dynamics of the Atmosphere*, Quarterly Journal of the Royal Meteorological Society, v. 113, 735-757, 1987

[12] M. J. P. Cullen: Private communication

[13] R. Daley: *Atmospheric Data Analysis*, Cambridge University Press, 1991

[14] M. Dixon and I. Roulstone: *Controlling Imbalance Within VAR by Use of a Weak Constraint*, Joint Centre for Mesoscale Meteorology Report No. 143, Met Office UK, August 2003

[15] M. J. Forray: *Variational Calculus in Science and Engineering*, McGraw-Hill, 1968

[16] P. Gauthier and J. Thépaut: *Impact of the Digital Filter as a Weak Constraint in the Preoperational 4d Var Assimilation System of Météo-France*, Monthly Weather Review, v. 129, 2089-2102

[17] R. Giering and T. Kaminski: *Recipes for Adjoint Code Construction*, Max-Planck-Institut für Meteorologie, Report No. 212, August 1996

[18] P. E. Gill, W. Murray and M. H. Wright: *Practical Optimization*, Academic Press, 1981

[19] H. Goldstein: *Classical Mechanics*, Addison-Wesley Press, 1951

[20] E. Hairer, C. Lubich and G. Wanner: *Geometric Numerical Integration Illustrated by the Störmer/Verlet Method*, Acta Numerica, 1-51, 2003

[21] W. Huang and B. Leimkuhler: *The Adaptive Verlet Method*, SIAM Journal of Scientific Computing, v. 18, no. 1, 239-256, January 1997.

[22] C. Johnson: *Information Content of Observations in Variational Data Assimilation*, PhD thesis, Department of Meteorology, University of Reading, September 2003

[23] T. W. B. Kibble: *Classical Mechanics*, McGraw-Hill Publishing Company, 1966

[24] A. Lawless: *Development of Linear Models for Data Assimilation in Numerical Weather Prediction*, PhD thesis, Dept of Mathematics, University of Reading, May 2001

[25] F. Le Dimet and O. Talagrand: *Variational Algorithms for Analysis and Assimilation of Meteorological Observations: Theoretical Aspects*, Tellus, v. 38A, 97-110, 1986

[26] J. W. Leech: *Classical Mechanics*, Chapman and Hall, 1965

[27] B. Leimkuhler: *An Efficient Multiple Time-Scale Reversible Integrator for the Gravitational N-Body Problem*, Applied Numerical Mathematics, v. 43, 175-190

[28] B. Leimkuhler: *Reversible Adaptive Regularization: Perturbed Kepler Motion and Classical Atomic Trajectories*, Philosophical Transactions of the Royal Society of London, v. A 357, 1101-1134

[29] Y. Li, I. M. Navon, P. Courtier and P. Gauthier: *Variational Data Assimilation with a Semi-Lagrangian Semi-implicit Global Shallow-Water Equation Model and Its Adjoint*, Monthly Weather Review, v. 121, 1759-1769

[30] Y. Li, I. Navon, W. Yang, X. Zou, J. Bates, S. Moorthi, and R. Higgins: *Four-Dimensional Variational Data Assimilation Experiments With a Multilevel Semi-Lagrangian Semi-Implicit General Circulation Model*, Monthly Weather Review, v. 122, 966-983

[31] A. C. Lorenc: *Analysis Methods for Numerical Weather Prediction*, Quarterly Journal of the Royal Meteorological Society, v. 112, 1177-1194, 1986

[32] A. C. Lorenc, S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne and F. W. Saunders,: *The Met. Office Global Three-Dimensional Variational Data Assimilation Scheme*, Quarterly Journal of the Royal Meteorological Society, v. 126, 2991-3012, 2000

[33] E. N. Lorenz: *Deterministic Nonperiodic Flow*, Journal of the Atmospheric Sciences, v. 20, 130-141, 1963

[34] P. Lynch: *Hamiltonian Methods for Geophysical Fluid Dynamics: An Introduction*, IMA University of Minnesota, Preprint 1838, March 2002

[35] P. Lynch: *The Slow Equations*, Quarterly Journal of the Royal Meteorological Society, v. 114, 201-219

[36] D. J. C. Mackay: *1B Dynamics: Kepler's Ellipses and Gravitational Slingshots*, Lecture notes, University of Cambridge, UK, http://www.inference.phy.cam.ac.uk/teaching/dynamics/tex/orbits.pdf, August 2006

[37] M. E. McIntyre and I. Roulstone: *Are There Higher-Accuracy Analogues of Semigeostrophic Theory?*, Large-scale Atmosphere-Ocean Dynamics II: Geometric Methods and Models, ed I. Roulstone and J. Norbury, Cambridge Universtiy Press, 2001

[38] Met Office UK: *Operational Numerical Modelling*, http://www.metoffice.com/research/nwp/numerical/operational/index.html, September 2006

[39] N. K. Nichols: *Treating Model Error in 3-D and 4-D Data Assimilation*, Data Assimilation for the Earth System, Kluwer Academic, 127-135, 2003

[40] N. K. Nichols: *Data Assimilation: Aims and Basic Concepts*, Proceedings of the NATO Advanced Study Institute on Data Assimilation for the Earth System, Maratea, Italy, June 2002

[41] The Nine Planets: *Appendix 1a: Solar System Data*, http://www.nineplanets.org/data.html, September 2006

[42] R. Pogge: *An Introduction to Solar System Astronomy: Lecture 19 Orbits*, Lecture notes, Ohio State University, http://www-astro.mps.ohio-state.edu/ pogge/Ast161/Unit4/orbits.html, August 2006

[43] R. Salmon: *Hamiltonian Fluid Mechanics*, Annual Review of Fluid Mechanics, v. 20, 225-256, 1988

[44] J. M. Sanz-Serna and M. P. Calvo: *Numerical Hamiltonian Problems*, Chapman and Hall, 1994

[45] Y. Sasaki: *Some Basic Formalisms in Numerical Variational Analysis*, Monthly Weather Review, v. 98, no. 12, 875-883

[46] D. F. Shanno and K. H. Phua: *CONMIN code*, Journal of ACM Transactions on Mathematical Software, Guide to Available Mathematical Software, National Institute of Standards and Technology, http://gams.nist.gov/serve.cgi/Module/TOMS/500/8535, August 2006

[47] R. D. Skeel and J. J. Biesiadecki: *Symplectic Integration with Variable Stepsize*, Annals of Numerical Mathematics, v. 1, 1-9, 1994

[48] V. Szebehely: *Stability in Dynamical Astronomy*, Proceedings of the National Academy of Sciences, v. 75, no. 12, 5743-5745, December 1978

[49] V. Szebehely and R. McKenzie: *Stability of the Sun-Earth-Moon system*, The Astronomical Journal, v. 82, no. 4, 303-305, 1977

[50] J. Thépaut and P. Courtier: *Four-Dimensional Variational Data Assimilation Using the Adjoint of a Multilevel Primitive Equation Model*, Quarterly Journal of the Royal Meteorological Society, v. 117, 1225-1254, 1991

[51] T. Wee and Y. Kuo: *Impact of a Digital Filter as a Weak Constraint in MM5 4d Var: An Observing System Simulation Experiment*, Monthly Weather Review, v. 132, 543-559

[52] D. Zupanski: *A General Weak Constraint Applicable to Operational 4d Var Data Assimilation Systems*, Monthly Weather Review, v. 125, 2274-2292

Appendix A

The Linear Three Body Equations

The linearisation of the nonlinear three body model equations is given by

$$
\begin{aligned}
\delta P_1^{n+\frac{1}{2}} &= \delta P_1^n + \frac{h}{2}((-\alpha_{12}^n + \beta_{X_{12}}^n - \alpha_{31}^n + \beta_{X_{31}}^n)\delta Q_1^n + (\gamma_{12}^n + \gamma_{31}^n)\delta Q_2^n \\
&\quad +(\alpha_{12}^n - \beta_{X_{12}}^n)\delta Q_3^n + (-\gamma_{12}^n)\delta Q_4^n \\
&\quad +(\alpha_{31}^n - \beta_{X_{31}}^n)\delta Q_5^n + (-\gamma_{31}^n)\delta Q_6^n) \qquad \text{(A.1)} \\
\delta P_2^{n+\frac{1}{2}} &= \delta P_2^n + \frac{h}{2}((\gamma_{12}^n + \gamma_{31}^n)\delta Q_1^n + (-\alpha_{12}^n + \beta_{Y_{12}}^n - \alpha_{31}^n + \beta_{Y_{31}}^n)\delta Q_2^n \\
&\quad +(-\gamma_{12}^n)\delta Q_3^n + (\alpha_{12}^n - \beta_{Y_{12}}^n)\delta Q_4^n \\
&\quad +(-\gamma_{31}^n)\delta Q_5^n + (\alpha_{31}^n - \beta_{Y_{31}}^n)\delta Q_6^n) \qquad \text{(A.2)} \\
\delta P_3^{n+\frac{1}{2}} &= \delta P_3^n + \frac{h}{2}((\alpha_{12}^n - \beta_{X_{12}}^n)\delta Q_1^n + (-\gamma_{12}^n)\delta Q_2^n \\
&\quad +(-\alpha_{12}^n + \beta_{X_{12}}^n - \alpha_{23}^n + \beta_{X_{23}}^n)\delta Q_3^n + (\gamma_{12}^n + \gamma_{23}^n)\delta Q_4^n \\
&\quad +(\alpha_{23}^n - \beta_{X_{23}}^n)\delta Q_5^n + (-\gamma_{23}^n)\delta Q_6^n) \qquad \text{(A.3)} \\
\delta P_4^{n+\frac{1}{2}} &= \delta P_4^n + \frac{h}{2}((-\gamma_{12}^n)\delta Q_1^n + (\alpha_{12}^n - \beta_{Y_{12}}^n)\delta Q_2^n \\
&\quad +(\gamma_{12}^n + \gamma_{23}^n)\delta Q_3^n + (-\alpha_{12}^n + \beta_{Y_{12}}^n - \alpha_{23}^n + \beta_{Y_{23}}^n)\delta Q_4^n \\
&\quad +(-\gamma_{23}^n)\delta Q_5^n + (\alpha_{23}^n - \beta_{Y_{23}}^n)\delta Q_6^n) \qquad \text{(A.4)}
\end{aligned}
$$

$$\begin{aligned}
\delta P_5^{n+\frac{1}{2}} &= \delta P_5^n + \frac{h}{2}((\alpha_{31}^n - \beta_{X_{31}}^n)\delta Q_1^n + (-\gamma_{31}^n)\delta Q_2^n \\
&\quad +(\alpha_{23}^n - \beta_{X_{23}}^n)\delta Q_3^n + (-\gamma_{23}^n)\delta Q_4^n + (-\alpha_{23}^n + \beta_{X_{23}}^n - \alpha_{31}^n \\
&\quad +\beta_{X_{31}}^n)\delta Q_5^n + (\gamma_{23}^n + \gamma_{31}^n)\delta Q_6^n) \qquad \text{(A.5)} \\
\delta P_6^{n+\frac{1}{2}} &= \delta P_6^n + \frac{h}{2}((-\gamma_{31}^n)\delta Q_1^n + (\alpha_{31}^n - \beta_{Y_{31}}^n)\delta Q_2^n \\
&\quad +(-\gamma_{23}^n)\delta Q_3^n + (\alpha_{23}^n - \beta_{Y_{23}}^n)\delta Q_4^n \\
&\quad +(\gamma_{23}^n + \gamma_{31}^n)\delta Q_5^n + (-\alpha_{23}^n + \beta_{Y_{23}}^n - \alpha_{31}^n + \beta_{Y_{31}}^n)\delta Q_6^n) \qquad \text{(A.6)}
\end{aligned}$$

$$\begin{aligned}
\delta Q_1^{n+1} &= \delta Q_1^n + \frac{h}{m_1}\delta P_1^{n+\frac{1}{2}} \qquad \text{(A.7)} \\
\delta Q_2^{n+1} &= \delta Q_2^n + \frac{h}{m_1}\delta P_2^{n+\frac{1}{2}} \qquad \text{(A.8)} \\
\delta Q_3^{n+1} &= \delta Q_3^n + \frac{h}{m_1}\delta P_3^{n+\frac{1}{2}} \qquad \text{(A.9)} \\
\delta Q_4^{n+1} &= \delta Q_4^n + \frac{h}{m_1}\delta P_4^{n+\frac{1}{2}} \qquad \text{(A.10)} \\
\delta Q_5^{n+1} &= \delta Q_5^n + \frac{h}{m_1}\delta P_5^{n+\frac{1}{2}} \qquad \text{(A.11)} \\
\delta Q_6^{n+1} &= \delta Q_6^n + \frac{h}{m_1}\delta P_6^{n+\frac{1}{2}} \qquad \text{(A.12)}
\end{aligned}$$

$$\begin{aligned}
\delta P_1^{n+1} &= \delta P_1^{n+\frac{1}{2}} + \frac{h}{2}((-\alpha_{12}^{n+1} + \beta_{X_{12}}^{n+1} - \alpha_{31}^{n+1} + \beta_{X_{31}}^{n+1})\delta Q_1^{n+1} \\
&\quad +(\gamma_{12}^{n+1} + \gamma_{31}^{n+1})\delta Q_2^{n+1} \\
&\quad +(\alpha_{12}^{n+1} - \beta_{X_{12}}^{n+1})\delta Q_3^{n+1} + (-\gamma_{12}^{n+1})\delta Q_4^{n+1} \\
&\quad +(\alpha_{31}^{n+1} - \beta_{X_{31}}^{n+1})\delta Q_5^{n+1} + (-\gamma_{31}^{n+1})\delta Q_6^{n+1}) \qquad \text{(A.13)} \\
\delta P_2^{n+1} &= \delta P_2^{n+\frac{1}{2}} + \frac{h}{2}((\gamma_{12}^{n+1} + \gamma_{31}^{n+1})\delta Q_1^{n+1} \\
&\quad +(-\alpha_{12}^{n+1} + \beta_{Y_{12}}^{n+1} - \alpha_{31}^{n+1} + \beta_{Y_{31}}^{n+1})\delta Q_2^{n+1} \\
&\quad +(-\gamma_{12}^{n+1})\delta Q_3^{n+1} + (\alpha_{12}^{n+1} - \beta_{Y_{12}}^{n+1})\delta Q_4^{n+1} \\
&\quad +(-\gamma_{31}^{n+1})\delta Q_5^{n+1} + (\alpha_{31}^{n+1} - \beta_{Y_{31}}^{n+1})\delta Q_6^{n+1}) \qquad \text{(A.14)}
\end{aligned}$$

$$\begin{aligned}\delta P_3^{n+1} &= \delta P_3^{n+\frac{1}{2}} + \frac{h}{2}((\alpha_{12}^{n+1} - \beta_{X_{12}}^{n+1})\delta Q_1^{n+1} + (-\gamma_{12}^{n+1})\delta Q_2^{n+1} \\ &+(-\alpha_{12}^{n+1} + \beta_{X_{12}}^{n+1} - \alpha_{23}^{n+1} + \beta_{X_{23}}^{n+1})\delta Q_3^{n+1} + (\gamma_{12}^{n+1} + \gamma_{23}^{n+1})\delta Q_4^{n+1} \\ &+(\alpha_{23}^{n+1} - \beta_{X_{23}}^{n+1})\delta Q_5^{n+1} + (-\gamma_{23}^{n+1})\delta Q_6^{n+1})\end{aligned}\tag{A.15}$$

$$\begin{aligned}\delta P_4^{n+1} &= \delta P_4^{n+\frac{1}{2}} + \frac{h}{2}((-\gamma_{12}^{n+1})\delta Q_1^{n+1} + (\alpha_{12}^{n+1} - \beta_{Y_{12}}^{n+1})\delta Q_2^{n+1} \\ &+(\gamma_{12}^{n+1} + \gamma_{23}^{n+1})\delta Q_3^{n+1} + (-\alpha_{12}^{n+1} + \beta_{Y_{12}}^{n+1} - \alpha_{23}^{n+1} + \beta_{Y_{23}}^{n+1})\delta Q_4^{n+1} \\ &+(-\gamma_{23}^{n+1})\delta Q_5^{n+1} + (\alpha_{23}^{n+1} - \beta_{Y_{23}}^{n+1})\delta Q_6^{n+1})\end{aligned}\tag{A.16}$$

$$\begin{aligned}\delta P_5^{n+1} &= \delta P_5^{n+\frac{1}{2}} + \frac{h}{2}((\alpha_{31}^{n+1} - \beta_{X_{31}}^{n+1})\delta Q_1^{n+1} + (-\gamma_{31}^{n+1})\delta Q_2^{n+1} \\ &+(\alpha_{23}^{n+1} - \beta_{X_{23}}^{n+1})\delta Q_3^{n+1} + (-\gamma_{23}^{n+1})\delta Q_4^{n+1} \\ &+(-\alpha_{23}^{n+1} + \beta_{X_{23}}^{n+1} - \alpha_{31}^{n+1} + \beta_{X_{31}}^{n+1})\delta Q_5^{n+1} \\ &+(\gamma_{23}^{n+1} + \gamma_{31}^{n+1})\delta Q_6^{n+1})\end{aligned}\tag{A.17}$$

$$\begin{aligned}\delta P_6^{n+1} &= \delta P_6^{n+\frac{1}{2}} + \frac{h}{2}((-\gamma_{31}^{n+1})\delta Q_1^{n+1} + (\alpha_{31}^{n+1} - \beta_{Y_{31}}^{n+1})\delta Q_2^{n+1} \\ &+(-\gamma_{23}^{n+1})\delta Q_3^{n+1} + (\alpha_{23}^{n+1} - \beta_{Y_{23}}^{n+1})\delta Q_4^{n+1} \\ &+(\gamma_{23}^{n+1} + \gamma_{31}^{n+1})\delta Q_5^{n+1} \\ &+(-\alpha_{23}^{n+1} + \beta_{Y_{23}}^{n+1} - \alpha_{31}^{n+1} + \beta_{Y_{31}}^{n+1})\delta Q_6^{n+1}),\end{aligned}\tag{A.18}$$

where $\boldsymbol{\delta Q}^n$ is the vector of position at the nth timestep, $\boldsymbol{\delta P}^n$ the vector of momenta of the three bodies.

The coefficients in the above equations are given by

$$\alpha_{12}^n = \frac{m_1 m_2}{((Q_3^n - Q_1^n)^2 + (Q_4^n - Q_2^n)^2)^{\frac{3}{2}}}\tag{A.19}$$

$$\alpha_{23}^n = \frac{m_2 m_3}{((Q_5^n - Q_3^n)^2 + (Q_6^n - Q_4^n)^2)^{\frac{3}{2}}}\tag{A.20}$$

$$\alpha_{31}^n = \frac{m_3 m_1}{((Q_1^n - Q_5^n)^2 + (Q_2^n - Q_6^n)^2)^{\frac{3}{2}}}\tag{A.21}$$

$$\beta^n_{X_{12}} = \frac{3m_1m_2(Q^n_3 - Q^n_1)^2}{((Q^n_3 - Q^n_1)^2 + (Q^n_4 - Q^n_2)^2)^{\frac{3}{2}}} \tag{A.22}$$

$$\beta^n_{X_{23}} = \frac{3m_2m_3(Q^n_5 - Q^n_3)^2}{((Q^n_5 - Q^n_3)^2 + (Q^n_6 - Q^n_4)^2)^{\frac{3}{2}}} \tag{A.23}$$

$$\beta^n_{X_{31}} = \frac{3m_3m_1(Q^n_1 - Q^n_5)^2}{((Q^n_1 - Q^n_5)^2 + (Q^n_2 - Q^n_6)^2)^{\frac{3}{2}}} \tag{A.24}$$

$$\beta^n_{Y_{12}} = \frac{3m_1m_2(Q^n_4 - Q^n_2)^2}{((Q^n_3 - Q^n_1)^2 + (Q^n_4 - Q^n_2)^2)^{\frac{3}{2}}} \tag{A.25}$$

$$\beta^n_{Y_{23}} = \frac{3m_2m_3(Q^n_6 - Q^n_4)^2}{((Q^n_5 - Q^n_3)^2 + (Q^n_6 - Q^n_4)^2)^{\frac{3}{2}}} \tag{A.26}$$

$$\beta^n_{Y_{31}} = \frac{3m_3m_1(Q^n_2 - Q^n_6)^2}{((Q^n_1 - Q^n_5)^2 + (Q^n_2 - Q^n_6)^2)^{\frac{3}{2}}} \tag{A.27}$$

$$\gamma^n_{12} = \frac{3m_1m_2(Q^n_3 - Q^n_1)(Q^n_4 - Q^n_2)}{((Q^n_3 - Q^n_1)^2 + (Q^n_4 - Q^n_2)^2)^{\frac{3}{2}}} \tag{A.28}$$

$$\gamma^n_{23} = \frac{3m_2m_3(Q^n_5 - Q^n_3)(Q^n_6 - Q^n_4)}{((Q^n_5 - Q^n_3)^2 + (Q^n_6 - Q^n_4)^2)^{\frac{3}{2}}} \tag{A.29}$$

$$\gamma^n_{31} = \frac{3m_3m_1(Q^n_1 - Q^n_5)(Q^n_2 - Q^n_6)}{((Q^n_1 - Q^n_5)^2 + (Q^n_2 - Q^n_6)^2)^{\frac{3}{2}}}, \tag{A.30}$$

where the vectors $\boldsymbol{Q}^n$ and $\boldsymbol{P}^n$ are the linearisation states.

Appendix B

Calculation of S_{cr}

From [48] we have that

$$S_{cr} = \frac{f^2 g}{6\bar{m}^6} \tag{B.1}$$

where

$$\begin{aligned} f(x) &= m_1 m_3 + \frac{m_2 m_3}{1+x} + \frac{m_1 m_2}{x} && \text{(B.2)} \\ g(x) &= m_1 m_3 + m_2 m_3 (1+x)^2 + m_1 m_2 x^2, && \text{(B.3)} \end{aligned}$$

and $\bar{m}$ is the mean mass of the three bodies.

The value of x is such that it satisfies

$$(m_3+m_1)x^5+(3m_3+m_1)x^4+(3m_3+m_1)x^3-(3m_2+m_1)x^2-(3m_2+2m_1)x-(m_2+m_1) = 0. \tag{B.4}$$

The values of the masses for our problem are $m_1 = 1$, $m_2 = 0.1$ and $m_3 = 0.01$. By substituting these values into (B.4) and using the Newton method to iteratively solve the resulting polynomial, we find that, for this problem, $x = 1.1546$. Using this we can evaluate $f(x)$ and $g(x)$. We therefore calculate that the critical value to be

$$S_{cr} = 9.0567 \times 10^{-2} \tag{B.5}$$