A Hybrid Symplectic Integrator for Fitting Exoplanet TTVs and Candidate System Stability Analysis

Matthew Sylvester

A thesis submitted in partial fulfillment
of the BSc Honours in Physics and Astronomy

Physics & Astronomy
Bishop's University
Canada

April 26th 2024

Contents

List of Figures

1 Abstract

Exoplanetary systems are examples of the N-body problem in action. This project replicates an approach to fitting the transit timing variations of exoplanets in perturbed Keplerian orbits through the application of a symplectic integrator. TTVs, deviations in an exoplanet's expected orbit, offer a method to probe the dynamical interactions present in a planetary system. We implement an N-body integrator designed to evaluate the goodness of fit of numerically calculated transits to observations in order to determine candidate configurations. Traditional integration schemes suffer from rapid loss of accuracy and divergence from the true solution when applied to chaotic systems, or have prohibitively expensive computational requirements. HSI algorithms leverage the symplectic nature of the dominant Keplerian motion of bodies in the system to ensure long-term stability. When the N-body interactions become dominant, the energy conserving Keplerian approximation to the motion is no longer valid, so the symplectic advantage is lost. A higher order integrator is thus required to maintain good accuracy during close encounters. The result is a prescription which both maintains stability in the long term while keeping computations low and adapts to maintain short term accuracy when large interaction terms become a concern.

2 Introduction

In early February of 1990, astronomers Aleksander Wolszczan and Dale Frail at the Aricebo radio telescope were making observations of the pulsar PSR B1257+12 in the constellation of Virgo. Two short years later, these observations would become the basis of the discovery of the first planetary bodies orbiting a star other than our own [1]. This marked the birth of the field of exoplanet astronomy. More discoveries quickly followed. In 1995, a Jupiter mass planet was found around the sun-like star 51 Pegasi using the radial velocity method. In 2002, the first planet was discovered using the transit method [2].

To date, the existence of over 5000 [3] exoplanets have been confirmed using various observational methods. The systems they inhabit are diverse in both the characteristics of their orbits and host star (though our methods of measurement bias the detection of certain configurations), so there exists a rich potential for scientific inquiry in their characterization. What is inescapably common between all of the exoplanets we observe, though, is that their measurable effects are for the most part indirect and small in magnitude. It's for this reason that we must make clever use of all the tools at our disposal to aid in our characterization of exoplanetary systems.

2.1 Scope of Work

The work outlined here involves the replication of an efficient computational method of modelling planetary systems in order to perform analysis of transit timing variations. These variations provide enough insight on the dynamical interactions taking place in a planetary system to attempt to put constraints on its orbital parameters. The details of the methods used are elaborated in section 4, with background information provided in section 3 to help motivate the methods used.

We successfully constructed a symplectic integrator, as well as a pipeline to use it to analyze the goodness of fit of an exoplanetary system's orbital configuration by comparing model Transit Timing Variations (TTVs) to Kepler data. A hybrid integrator which makes use of a higher order Bulirsch-Stoer integrator to maintain accuracy during close encounters was also implemented. The hybrid scheme is used to test the stability of candidate configurations over long ranges while keeping compute time low.

Our integrator was built using the Julia programming language, and employed the use of the StaticArrays, LinearAlgebra, DelimitedFiles and Interpolations packages. Figures were generated using Julia's native Plots package and in LaTeX using TikZ.

3 Background

3.1 Orbital Dynamics

This section primarily follows the derivations found in Murray & Dermot's Solar System Dynamics [4].

Gravitation is the fundamental force that determines the motion of systems on the scale of planetary systems. A basic understanding of the principles which drive the motion of gravitationally bound bodies is the first step in developing a framework for extracting information from exoplanet observations.

Beginning from the basic governing equation for gravitationally interacting bodies,

$$\boldsymbol{F_1} = \frac{Gm_1m_2}{r^3}\mathbf{r} = m_1\ddot{\boldsymbol{r}}_\mathbf{1} \tag{3.1.1}$$

Here, $\boldsymbol{F_1}$ is the vector force acting on body 1, m_n is the mass of body n, and $\boldsymbol{r_1}$ is the vector position of the body from some inertial origin O. The vector $\boldsymbol{r} = r_2 - r_1$ defines the relative position of body 2 with respect to body 1. Equation 3.1.1 describes the motion of a body in a two-body system in its orbit around the centre of mass, or the barycentre.

Summing the forces on each object, we have

$$m_1\ddot{\boldsymbol{r}}_\mathbf{1} + m_2\ddot{\boldsymbol{r}}_\mathbf{2} = 0 \tag{3.1.2}$$

which, when integrated twice, gives

$$m_1\dot{\boldsymbol{r}}_\mathbf{1} + m_2\dot{\boldsymbol{r}}_\mathbf{2} = \boldsymbol{a} \quad \text{and} \quad m1\boldsymbol{r_1} + m_2\boldsymbol{r_2} = \boldsymbol{a}t + b \tag{3.1.3}$$

Where $\mathbf{a}$ and $\mathbf{b}$ are constant vectors. The barycentre is defined as $\boldsymbol{R} = \frac{m_1\boldsymbol{r_1}+m_2\boldsymbol{r_2}}{m_1+m_2}$. Substituting this into our equation above gives

$$\boldsymbol{R} = \frac{\boldsymbol{a}t + \boldsymbol{b}}{m_1 + m_2} \tag{3.1.4}$$

Which tells us that the velocity of the barycentre is either described by a constant vector $\boldsymbol{a}$ or $\hat{\mathbf{0}}$. This means that the barycentre is a valid candidate for our inertial point of reference.

Without loss of generality, we can consider the motion of one of the bodies with respect to the other. Setting $\boldsymbol{r} = \boldsymbol{r_2} - \boldsymbol{r_1}$

$$\frac{d^2\boldsymbol{r}}{dt^2} + \mu\frac{\boldsymbol{r}}{r^3} = 0 \tag{3.1.5}$$

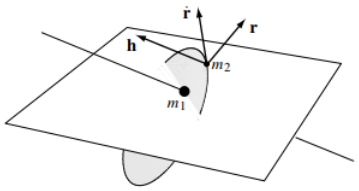

Figure 1: The orbital plane that confines the motion of the body m_2 about m_1 [4]

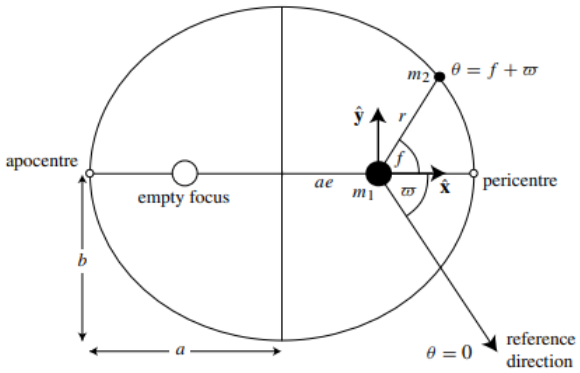

Figure 2: The full coordinate system used to describe orbital motion [4]

where $\mu = G(m_1 + m_2)$. Taking the vector product of $\mathbf{r}$ and equation 3.1.5 gives $\boldsymbol{r} \times \ddot{\boldsymbol{r}} = 0$, which integrates to $\boldsymbol{r} \times \dot{\boldsymbol{r}} = \boldsymbol{h}$, where $\boldsymbol{h}$ is a constant vector. This allows us to define a plane in which the motion of body 2 about body 1 will reside, with $\boldsymbol{h}$ as its normal.

We can now define a two dimensional polar coordinate system which lies in this plane in preparation of finding an equation to describe the orbital motion of body 2. Figures 1 and 2 show the basic geometry of the orbital plane. The use of an ellipse to describe the path of the orbiting object will be motivated soon.

In our new coordinate system, the position, velocity, and acceleration are written as

$$\boldsymbol{r} = r\hat{\boldsymbol{r}} \tag{3.1.6}$$

$$\dot{\boldsymbol{r}} = \dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}} \tag{3.1.7}$$

$$\ddot{\boldsymbol{r}} = (\ddot{r} - r\dot{\theta}^2)\hat{\boldsymbol{r}} + [\frac{1}{r}\frac{d}{dt}(r^2\dot{\theta})]\hat{\boldsymbol{\theta}} \tag{3.1.8}$$

Combining equations 3.1.5 and 3.1.8 gives us a differential equation in polar coordinates which encapsulates the force of gravitation experienced by m_2.

$$\ddot{r} - r\dot{\theta}^2 = -\frac{\mu}{r^2} \tag{3.1.9}$$

If we substitute our equation for $\dot{\boldsymbol{r}}$ into our equation for $\boldsymbol{h}$, we get $\boldsymbol{h} = r^2\dot{\theta}\hat{\mathbf{z}}$. In this non-inertial co-ordinate system, this vector represents a sort of pseudo rotational inertia. In the specific case of $m_1 >> m_2$, the motion of m_1 as it orbits around the barycentre of our two body system can be ignored and $\boldsymbol{h}$ represents the true rotational inertia of m_2.

We can make use of the constant $h = r^2\dot{\theta}$ to solve 3.1.9. After performing the substitution $u = 1/r$, and differentiating WRT time, our equation becomes

$$\frac{d^2u}{d\theta^2} + u = \frac{\mu}{h^2} \tag{3.1.10}$$

Which, after undoing our substitution, has the solution

$$r = \frac{p}{1 + e\cos(\theta - \bar{\omega})} \tag{3.1.11}$$

Where e is an amplitude for the variation in u, which is the eccentricity of the orbital ellipse, and $\bar{\omega}$ is the phase. Here, p is a stand-in for $h^2/G(m_1 + m_2)$.

This is the general equation for a set of curves called conic sections. The four possible conic sections are:

circle:	$e = 0$
ellipse:	$0 < e < 1$
parabola:	$e = 1$
hyperbola:	$e > 1$

While all of these conics have great relevance to orbital dynamics, we will be particularly concerned with bodies which have circular or elliptical orbits, with eccentricity less than 1. Parabolic and hyperbolic orbits are by definition unbounded and non-periodic, which makes it impossible to perform the sort of analysis we are concerned with. In fact, the majority of known planetary orbits have eccentricity $e << 1$, and are nearly circular. This is especially true of the orbits of bodies of particular interest for TTV measurements. Circular orbits

are expected for short-period planets as a result of their short tidal circularization timescale, and all but one of the currently known short period planets (<4.3 days) have eccentricities consistent with zero [5].

It's important to re-iterate that the derivation of this exact curve for the orbit of a body in a planetary system is valid only for the limited case of the two-body system. When there are more bodies in a planetary system, the path taken by the orbiting body will not be perfectly elliptical. This is why we must numerically integrate our orbits in the N-body case.

3.2 Keplerian Co-ordinates

The parameter p is given by

$$p = a(1 - e^2) \tag{3.2.1}$$

Where a is the semi-major axis, defined as the longest radius of the orbital ellipse. The relationship between a and the eccentricity e is

$$b^2 = a^2(1 - e^2) \tag{3.2.2}$$

Which lets us parameterize the equation for the radius at any given time as

$$r = \frac{a(1 + e^2)}{1 + ecos(\theta - \bar{\omega})} \tag{3.2.3}$$

The maxima and minima of the radial distance are called the *apoapsis* and *periapsis* respectively. The angle θ is called the true longitude. It is in reference to some arbitrary reference direction on the orbital plane. However, it can be convenient to define this reference direction in terms of the orbital parameters of the body. We introduce the *true anomaly*, defined as $f = \theta - \bar{\omega}$, which sets the reference direction as the direction which points towards the periapsis.

In order to relate our parameters directly to time, we define the mean anomaly

$$M = \frac{2\pi}{T}(t - \tau) \tag{3.2.4}$$

where τ is the *time of periapsis passage*, a constant which defines the absolute time at which the orbiting body passed . We also introduce the independent parameter the *eccentric anomaly* E, defined as the angle between the major axis of the ellipse and the radius from the centre to the intersection point on the circumscribed circle, or the circle that is centred on the centre of the ellipse with radius a.

The relationship between E, r, and e is

$$r = a(1 - e\cos E) \tag{3.2.5}$$

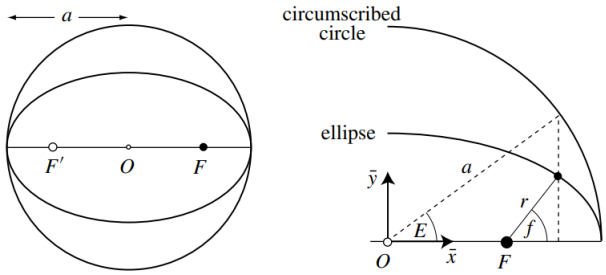

Figure 3: The circumscribed circle and the relationship between the eccentric anomaly and true anomaly f [4]

We must also define the magnitude of the angle made between the planet's orbital plane and the reference plane, as the orbital plane is not generally constant in any case but for two bodies. This angle is called the *inclination.* The point on the orbital path where the orbital and reference planes intersect is called the *ascending node,* and the angle between it and the reference direction is the *longitude of ascending node* Ω.

The longitude of ascending node is related to the argument of periapsis and the longitude of periapsis by

$$\bar{\omega} = \Omega + \omega \tag{3.2.6}$$

Seven independent parameters are required to describe the motion of a body in Keplerian orbital elements. They are as listed in table 1 (along with other relevant measurements)

3.3 Jacobi Co-ordinates

When describing the motion of multiple bodies using Keplerian orbital elements, some approximation is inevitable. The presence of additional gravitational forces means that, if we are only considering the central star when describing an orbiting body's Keplerian co-ordinates, there may be a significant error in the instantaneous values.
In an attempt to remedy this, Jacobi coordinates can be used as an intermediary when determining a body's Keplerian co-ordinates. Instead of referencing a central star, Jacobi coordinates place bodies relative to the barycentre of all bodies with lower indices. The justification for the use of this co-ordinate system is fairly intuitive. If one approximates the masses of every other planet as a shell located at their respective orbits, the gravitational influence of the interior shells can be approximated as point sources at the centre of their orbits, while the influence of the exterior shells will be zero [6] . There will still be some

a	Semi-major axis
e	Eccentricity
i	Inclination
ω	Argument of periapsis
τ	Time of periapsis passage
E	Eccentric Anomaly
f	True anomaly
M	Mean anomaly
$\bar{\omega} = \omega + \Omega$	Longitude of periapsis
$\theta = f + \bar{\omega}$	True longitude
$\lambda = M + \bar{\omega}$	Mean longitude

Table 1: Keplerian orbital elements and their symbols [4]

approximation of the orbital elements when using Jacobi coordinates depending on the phases of the other planets, but on average, it is a good approximation. The Jacobi coordinate of the i-th body $\vec{r'}$ relative to the $i-1$-th barycentre $\vec{R_{i-1}}$ is

$$\vec{r'} = \vec{r_i} - \vec{R_{i-1}} \tag{3.3.1}$$

Where $\vec{R_i}$ is defined as

$$\vec{R_i} = \frac{1}{M_i} \sum_{j=0}^{i} m_j \vec{r_j} \tag{3.3.2}$$

$$M_i = \sum_{j=0}^{i} m_j \tag{3.3.3}$$

The velocities transform in the same way, with

$$\vec{v'} = \vec{v_i} - \vec{V_{i-1}} \tag{3.3.4}$$

Where $\vec{V_i}$ is defined as

$$\vec{V_i} = \frac{1}{M_i} \sum_{j=0}^{i} m_j \vec{v_j} \tag{3.3.5}$$

$$\tag{3.3.6}$$

3.4 The Hamiltonian Formulation

For reasons that will become clear in the subsequent sections, when we scale up our problem to include the interactions from the other bodies in the system and are forced to switch to numerical methods, it is helpful to first describe the system using Hamiltonian formalism. This will allow us to construct our integrator in a manner that it has the property of being *symplectic*.

Symplectic integrators are desirable because they preserve some properties of a Hamiltonian system. These integrators cut down on computational complexity by allowing us to constrain the phase-space evolution of the system under the Hamiltonian flow.

In Hamiltonian formalism, the canonical phase space co-ordinates are the position $\mathbf{q}$ and the momentum $\mathbf{p}$. These co-ordinates are independent, and together form the basis of the phase space, upon which we can completely describe the state of a Hamiltonian system. The equations which define the rate of change of these parameters are

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \tag{3.4.1}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \tag{3.4.2}$$

Where H is the Hamiltonian, defined as the sum of the kinetic and potential energies. In the gravitational N-body system, the Hamiltonian is

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} - G\sum_{i=1}^{N} m_i \sum_{j=i+1}^{N} \frac{m_j}{r_{ij}} \tag{3.4.3}$$

The rate of change of any parameter x can be expressed as

$$\frac{dx}{dt} = \sum_{i=1}^{N} \left(\frac{\partial x}{\partial q_i}\frac{dq}{dt} + \frac{\partial x}{\partial p_i}\frac{dp_i}{dt}\right) \tag{3.4.4}$$

$$= \sum_{i=1}^{N} \left(\frac{\partial x}{\partial q_i}\frac{dH}{dp_i} + \frac{\partial x}{\partial p_i}\frac{dH}{dq_i}\right) \tag{3.4.5}$$

We can call this the operator F, such that

$$\frac{dx}{dt} = Fx \tag{3.4.6}$$

Which has the general solution with timestep τ

$$x(t) = e^{(\tau F)}x(t-\tau) \tag{3.4.7}$$

Where τ is the timestep, and $x(t-\tau)$ is the value of x at the previous timestep. If we apply two operators (Say A & B), similar to F but not necessarily involving the same H, on x one after the other, we get

$$x(t) = e^{(\tau A)} e^{(\tau B)} x(t-\tau) \tag{3.4.8}$$

Which has the expansion

$$e^{\tau(A+B)} = 1 + \tau(A+B) + O(\tau^2) \tag{3.4.9}$$

Which gives us a first-order prescription for our integrator when we split the Hamiltonian into separate parts. This allows us to separate the dominant Keplerian terms and the perturbations of the n-body interactions [7]. This can be done using many different methods. Wisdom & Holman [8] suggest a separation using Jacobi Co-ordinates, into the Keplerian (H_{Kepler}) and interaction ($H_{interaction}$) terms. The result is a scheme which follows an alternating sequence, first evolving the system using the dominant Keplerian hamiltonian, followed by periodic kicks from the perturbation terms. The Wisdom & Holman split is of the form

$$H_{Kepler} = \sum_{i=1}^{n-1} \frac{p_i'^2}{2m_i'} - \frac{G m_i m_0}{r_i'} \tag{3.4.10}$$

$$H_{Interaction} = \sum_{i=1}^{n-1} G m_i m_0 (\frac{1}{r_i'} - \frac{1}{r_{i0}}) - \sum_{0<i<j} \frac{G m_i m_j}{r_{ij}} \tag{3.4.11}$$

Where primes denote Jacobi coordinates, $\vec{p_i'} = m_i' \vec{v_i'}$, and $m_i' = \sum_{j=0}^{i-1} m_{i-1} \frac{m_i}{\sum_{j=0}^{i-1} m_{i-1}}$.

Alternatively, Levison et Al [9] provide a splitting based on *democratic heliocentric co-ordinates*. This splitting avoids the requirement to keep track of both the Cartesian and Jacobian co-ordinates throughout the integration. In the democratic heliocentric scheme, positions are measured relative to the central star, with velocities measured relative to the barycentre of the entire system. In democratic heliocentric co-ordinates, the Hamiltonian splitting is

$$H_{Kepler} = \sum_{i=1}^{N} (\frac{p_i^2}{2mi} - \frac{G m_\odot m_i}{r_{i\odot}}) \tag{3.4.12}$$

$$H_{Interaction} = -G \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{m_i m_j}{r_{ij}} \tag{3.4.13}$$

$$H_{Sun} = \frac{1}{2m_\odot} (\sum_{i=1}^{N} \vec{p_i})^2 \tag{3.4.14}$$

In this splitting, H_{Kepler} and $H_{Interaction}$ serve the same function as in the previous splitting, describing the contributions of the keplerian and interaction

terms to the hamiltonian respectively. The final term H_{Sun} has to do with the reflex motion of the sun, and its appearance is the result of the fact that we have fixed the star at the centre of our co-ordinate system [7].

3.5 Numerical Integration

In general, numerical integrators attempt to find the solution to a differential equation over a given bound. One of the simplest numerical differential equations solvers is the Euler method [10]. Beginning with the limit definition of the derivative,

$$f' = \lim_{\tau \to 0} \frac{f(t+\tau) - f(t)}{\tau} \tag{3.5.1}$$

Which is Taylor expanded to the following

$$f(t+\tau) = f(t) + \tau f' t + \frac{\tau}{2} f''(t) + ... \tag{3.5.2}$$

Since τ is small, we now truncate the expansion, resulting in the first order approximation

$$f(t+\tau) \approx f(t) + \tau f'(t) \tag{3.5.3}$$

Which gives us an expression for the solution of f at time $t+\tau$. This truncation step is not exact when using finite values of τ, which is where the intrinsic error when using this class of integrator hides. In the usual notation for numerical steps, we have

$$y_{n+1} = y_n + hf(t_n, y_n) \tag{3.5.4}$$

where h is the timestep τ, and $f(t_n, y_n)$ is the derivative of y evaluated at timestep t_n. For a mechanical system, $y(t)$ is interpreted as the position, and $y'(t) = f$ as the velocity. If there are acceleration terms, though, we also have to introduce the second order equation

$$y''(t) = a(t, y) \tag{3.5.5}$$

How do we solve this using the scheme we developed earlier? by recognizing that the acceleration is equivalent to taking the derivative of the velocity $y'(t) = v(t)$. This allows us to reduce our second order differential equation to a series of coupled first order equations.

$$y'(t) = v(t) \tag{3.5.6}$$
$$v'(t) = a(v(t), y(t)) \tag{3.5.7}$$

Using expression 3.5.3, our system becomes

$$y(t+\tau) = y(t) + \tau v(t) \tag{3.5.8}$$
$$v(t+\tau) = v(t) + \tau a(v(t), y(t)) \tag{3.5.9}$$

In the step notation, this is

$$y_{n+1} = y(n) + hv(n) \tag{3.5.10}$$
$$v_{n+1} = v(n) + ha(n) \tag{3.5.11}$$

The following is a plot of a simple parabolic path, as solved by an Euler integrator, compared to the exact solution.

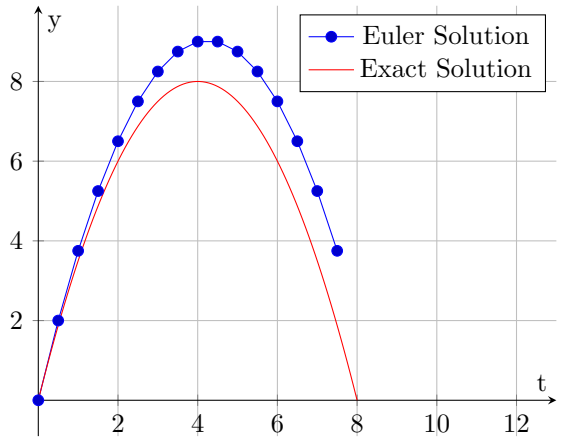

Figure 4: The evolution of an Euler integrator's solution compared to the true analytical solution with an uncharactaristically large timestep.

We can immediately see an issue with the Euler integrator; there is a secular trend in the error, as it is constantly compounded on the same side of the solution. Higher order integrators which follow the same principle as Euler's method exist. These methods increase the number of terms in the Taylor expansion of the derivative which are preserved, and thus have better accuracy. For example, the commonly used Runge Kutta method *RK4* [10], which preserves the 4th order term, is written as

$$y_{n+1} = y(t+h) \approx y(t) + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{3.5.12}$$

$$k_i = f(t, y(t), k_{i-1}) \tag{3.5.13}$$

The higher order terms in the expansion are encapsulated in k_i. They are approximated by successively performing Euler steps on fractions of the time

between t and $t + \tau$. This method is a good general purpose integrator which can traverse almost any solution without grave concern for fast accumulation of secular error or strange behaviour near singularities. When considering a problem in which the shape of the solution is fairly well known, however, we can do better.

3.6 Symplectic Integrators

Symplectic integrators are a class of integrator which take advantage of conservative systems.

Typically, when attempting to increase the accuracy of a conventional integrator, we want to reduce the error associated with evolving our system over a single time step. With symplectic integrators, we instead look at the long-term behaviour of our solution, accepting a somewhat larger step-by-step error in exchange for a solution which preserves a property of a conservative system. In our case, this is the Hamiltonian, which can be interpreted as the sum of the kinetic and potential energies $H = E_{tot} = T + U$. This proves particularly useful when modelling TTVs. A proportionality exists between the energy of an orbiting object and its period $P \propto E^{-3/2}$. By taking advantage of an integrator which by design conserves the energy of our system, we can avoid unnecessary contributions to our TTV model measurements due to error in the period caused by drift in the energy error.

When a system is plotted in phase space, we can define the concept of *flow* [11]. The flow of a system is a mapping which allows us to associate any point of a solution $y(t)$ with the with some initial value point in phase space $y(t = 0)$. The goal of a numerical integrator is to provide a map which replicates the phase space flow of the exact solution. Figure 5 illustrates the flow of a set of points in phase space when evolved under the exact flow of some arbitrary system and under the approximate numerical flow.

When we are dealing with a time independent Hamiltonian system, that is, one in which energy is conserved, we can take advantage of some properties of the phase space flow to inform the construction of our integrator. Liouville's theorem tells us that the Hamiltonian mapping from phase space co-ordinates (q, p) at $t = 0$ to (q', p') at $t = \tau$ along the solution is a canonical (symplectic) transformation [12], that is, it preserves the bivector

$$dp \wedge dq = dp' \wedge dq' \tag{3.6.1}$$

Which is equivalent to saying that phase space volume is conserved [13]. The determinant of the Jacobian of a mapping gives us the ratio between the phase space volume in our original co-ordinates (p, q) and the transformed co-ordinates (p', q'). So, to show that an integrator is symplectic is equivalent to demonstrating that the determinant of its Jacobian is unity. This is useful to us, because

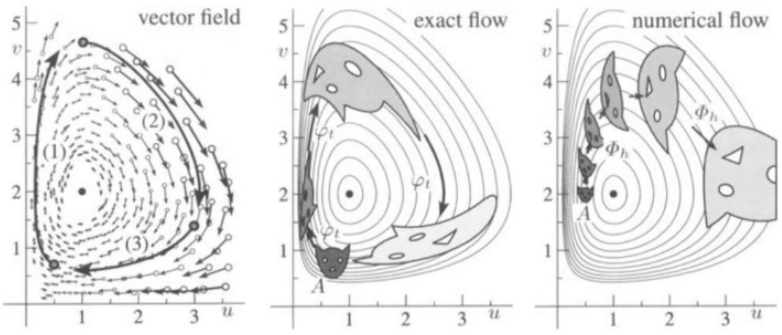

Figure 5: A phase space vector field, its exact flow, and an example of numerical flow. [11]

canonical transformations will evolve a point in phase space along the line of equal energy. If we can find a numerical map which satisfies the conditions of a symplectic transformation, we can be sure that our solution will follow along that line of equal energy as it is advanced along each time step.

With the knowledge that a given numerical scheme satisfies the symplectic condition, we can be somewhat more liberal when choosing the order of our integrator, which we will use to drastically reduce the number of computations required to integrate a solution to a given accuracy. Figure 6 illustrates the phase space evolution of an area of points of the harmonic oscillator using a low-order symplectic integrator as compared to the exact solution. While there is some significant step-by-step error in p and q, this error is symmetric, and the solution oscillates about a loop in phase space defined by a path of equal energy.

3.7 the Hybrid Modification

When close encounters occur, the perturbed Keplerian approximation to the solution of our system no longer works. A close encounter is a single event, so we cannot expect its effect to average out with other interactions in the long-term. This means we must consider the evolution of our system along more dimensions in phase space; those being the relative position and momenta of all the individual bodies. While it is simple to design an integrator which has the symplectic property for the Keplerian positions and momenta, as we can pick our (q_i, p_i) to be orthogonal, This is no longer the case when we must consider all $6n$ dimensions in the general n-body system. Numerical symplectic integrators tend to lose the symplectic property when they evolve a system over non-orthogonal phase space co-ordinates. This is unavoidable when considering more than 3 position and 3 momentum co-ordinates. This is because, when the solution is evolved along two non-orthogonal co-ordinates simultaneously, the

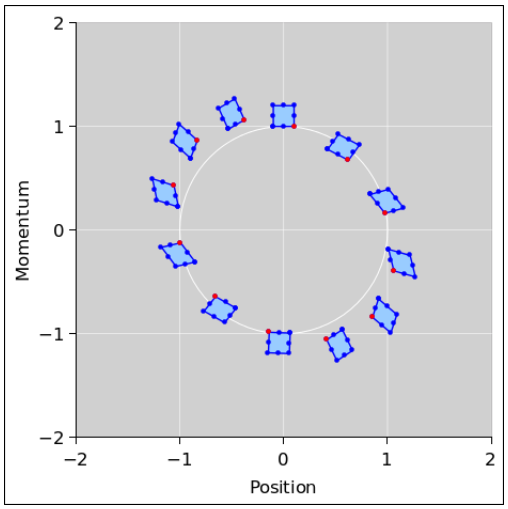

Figure 6: The phase space evolution of a simple harmonic oscillator under a symplectic mapping. An ensemble of oscillators shows Liouville's theorem in action; the area enclosed by the oscillators is constant.

oscillation about the path of equal energy is interrupted.

Without the symplectic property, during close encounters, all we are left with is a low order and unoptimised integrator. We could remedy this by adjusting the time step during close encounters, but this has two downsides.

1. The introduction of a variable time step further interrupts the oscillation of the solution about the path of equal energy. A smaller time step will reduce the magnitude of the oscillation, but we cannot be sure that the phase will be such that our solution lies on the correct path when the time step is decreased. Mathematically, the issue is that the Jacobian is no longer generally unity when we allow the time step τ to be a function of t.

2. An integrator designed to be symplectic may not take advantage of some other property of the problem to provide a lower step-by-step error, which is the strategy we must default to when we are no longer sure of the approximate total phase space path.

It's for these reasons that we make the switch to an optimized higher order integrator when close encounters occur for the long term integrations.

3.8 Gravitational Dominance & the Hill Radius

It is important to know when bodies in our integrator will be in the symplectic regime versus when they must be passed to the higher order integrator. For

Method of Detection	Number of Planets
Astrometry	3
Imaging	69
Radial velocity	1089
Transit	4171
Transit timing variations	29
Eclipse timing variations	17
Microlensing	214
Pulsar timing variations	8
Pulsation timing variations	2
Orbital brightness modulations	9
Disk kinematics	1

Table 2: Number of exoplanets by method of detection [3]

this, we turn to the concept of the *sphere of influence*, and how it relates to the gravitational dominance of a planetary body [14].

Most commonly, the Hill sphere is used to define the sphere of gravitational influence of a body. The radius of the hill sphere stretches from a body to the L_1 and L_2 Lagrange points. More exactly, for an approximately circular orbit, the radius of the hill sphere is defined as

$$R_H = (\frac{M}{M_\odot})^{1/3} a \tag{3.8.1}$$

where $M/M_\odot$ is the mass ratio between the orbiting body and the central star, and a is the semimajor axis. The ratio between the forces imparted upon an orbiting object by a planet and its star at the extrema of the hill sphere is always 2/3, regardless of the masses involved [15].

3.9 Exoplanet Detections & The Transit Method

Various methods have been employed to detect and measure the properties of exoplanets. Table 2 lists the number of detections associated with different methods as of 2024.

The first method of detection was the "wobble" method, which used the periodic red and blue shifting caused by the orbit of a massive planet around its host star. Systems which contain close-orbiting planets with masses within a

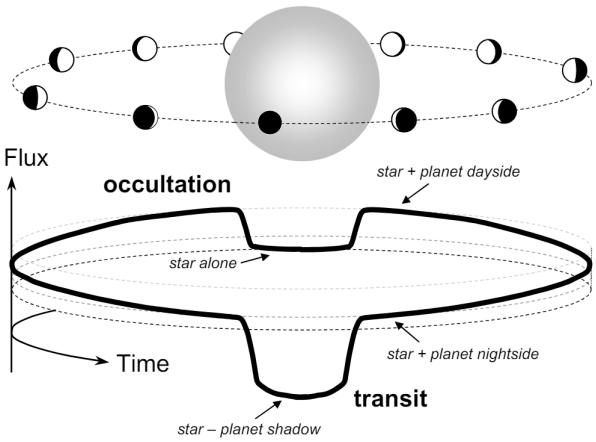

Figure 7: Diagram of the light curve of a transiting exoplanet [16]

few orders of magnitude of their host star are ideal for this method [3].

The most successful means of detection by good margin is the transit method. Detecting exoplanets by transits involves watching for a dip in relative flux as the planet passes between the earth and the host star. It is possible to constrain periods extremely well using this method, with error on the scale of seconds to minutes, as transits are short, distinct events. The transit method does require the system to be edge-on and co-planar. Luckily, the mechanics of galaxy formation and the conservation of angular momentum have driven most planetary systems within the milky way to evolve to be close to co-planar with the galaxy, and our solar system. This means we have a good chance of finding candidates.

Figure 7 illustrates the anatomy of a transit. As a transiting exoplanet travels in its orbit, it periodically passes directly between its host star and the earth. By measuring the relative flux of the star over time, we can detect these periodic dimming events, and determine the time at which the transit occurs with accuracy on the order of seconds.

Considering only inferior conjunctions, where the planet is in front of the host star, transits in a two-body system will occur at the regular cadence

$$t_{i,n} = n * P_i + t_0 \tag{3.9.1}$$

where n refers to the n'th transit of the planet, and t_0 the time of the initial transit.

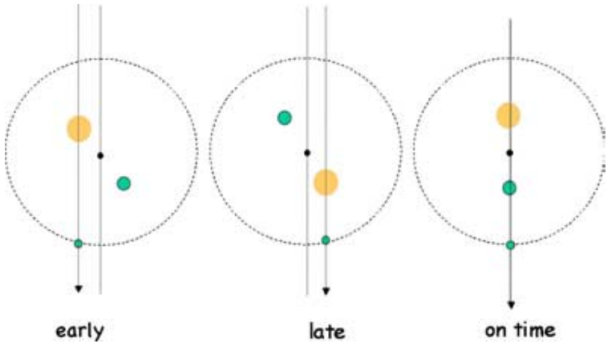

Figure 8: TTVs caused by reflex motion of host star due to massive planet. [20]

3.10 Transit Timing Variations

When solving a gravitational system containing only two bodies, we can be sure that their motion is perfectly encapsulated by the equations described in section 3.1, as there is one integral corresponding to each degree of freedom. When we involve multiple bodies, we can look to perturbations for constraints.

When looking at transit measurements, there are different parameters that can be examined for the effects of perturbations. Transit Timing Variations (TTV), Transit Duration Variations, and Transit Depth Variations are three of the major parameters often examined [17][18]. This work will be exclusively interested in examining the dynamical information embedded in TTVs. This method, combined with light curve fitting, has previously been used to constrain the parameters of exoplanetary systems [19].

To motivate the use of TTVs to characterize a system, we will consider the reduced problem of two non-interacting planets orbiting around a star. In our model system, the inner planet is massive enough to cause appreciable displacement of the sun-planet binary's barycentre, while the outer planet is much smaller and does not have this effect. Figure 8 is a representation of this model system. In this case, it can be shown that [20] the displacement of the star from the barycentre by the inner planet is given by

$$x_0 = -a_1\mu_1 sin(2\pi(t - t_0)/P_1) \tag{3.10.1}$$

Where μ is the reduced mass, and the transit time of the inner binary is t_0. This is known as the *reflex motion*. The timing deviation of the outer planet's transit due to this reflex motion is

$$\delta t_2 \approx \frac{x_0}{v_2 - v_0} \approx -\frac{P_2 a_1 \mu_1 \sin(2\pi(mP_2 - t_0)/P_1)}{2\pi a_2} \tag{3.10.2}$$

The standard deviation of this variation over many orbits is

$$\sigma_2 = \frac{P_2 a_1 \mu_1}{2^{3/2}\pi a_2} \tag{3.10.3}$$

It's important to note that these equations only apply to circular orbits, and only for the simplified problem of perfectly circular orbits and a non-interacting third body. While we have already discussed the lack of interaction effects on the TTVs of planets in approximately circular, eccentricities can change this. It is possible, albeit more complicated than the case demonstrating reflex motion, to show that the effect of an interacting planet with an eccentric orbit on the TTV of an inner planet-star binary is given by [20]

$$\delta t_1 = \frac{m_2}{2\pi(m_0 + m_1)} \frac{P_1^2}{P_2} (1 - e_2^2)^{-3/2} (f_2 - n_2(t - \tau_2) + e_2 \sin f_2) \tag{3.10.4}$$

This approximation works well in the limit that $r_1 << r_2$. TTVs are typically represented as a value $O-C$, or *Observed - Calculated*, a measure of the observed transit time minus the expected transit time if the orbit were perfectly circular for the n'th orbit $t_n = n * P + t_0$.

4 Methods

4.1 Units & Dimensions

The units fundamental units chosen for mass, time, and length were the earth mass $m_{\oplus}$, *year*, and AU. The use of these units inside the symplectic integrator helps to reduce round off floating point error by keeping most values within $\tilde{2}$ to 3 orders of magnitude of unity, though the interaction distance terms may become rather small at times. Before varying the initial parameters of our integrations, we translate the Cartesian co-ordinates into Keplerian orbital elements (E, P, e, ω, m), as this is a much more orthogonal dataset, which is preferable when attempting this sort of analysis. How the system acts is much more correlated with where we are in Keplerian parameter space.

While Keplerian orbital elements are preferable when traversing parameter space for our system's initial conditions, Cartesian co-ordinates are used for the integrator throughout a single integration. This is because the separation of variables in the differential equations governing the motion of the bodies simple in these co-ordinates. Rather than using Jacobi co-ordinates as Wisdom & Holman's original symplectic integrator, we used Democratic Heliocentric Co-ordinates. This results in a slightly more complicated Hamiltonian splitting, as described in section 3.4, but it avoids the requirement of re-calculating the Jacobi co-ordinates during each step, as the interaction portion of the Hamiltonian still required the usual Cartesian positions.

In order to limit the scope of the project, and to reduce the dimensions of the configuration space of our system, we considered only four of the orbital elements, while assuming the planets in the system lie only on the orbital plane. This allowed us to reduce our analysis to four elements per planet;

- Period P
- Eccentric Anomaly E
- Longitude of Pericentre ω
- Eccentricity e

These four elements, along with the mass of the body, allow us to completely describe the ellipse along which it travels, as well as its phase. The period and eccentricity together define the shape of the ellipse, as P is directly proportional to the semimajor axis, the longitude of pericentre its orientation in relation to a reference angle, and the eccentric anomaly the phase.

4.2 The Symplectic Integrator

The integration scheme developed for the symplectic portion of the integrator follows the method described by Wisdom & Holman. [8] Our scheme uses the Democratic Heliocentric splitting, however, so there is an extra step required during each iteration. The algorithm is as follows;

Algorithm 1 Symplectic Integrator Step

1: Evolve the system under H_C by shifting each body by $\tau \sum_i p_i/2m_\odot$ for $\tau/2$
2: Evolve the system under H_B by giving each body the perturbative kicks of all other bodies but the central star for $\tau/2$
3: Evolve the system under H_A by;
- Giving each body a symplectic Keplerian kick for τ
- Advancing the position of each body based on its momentum for τ

4: Evolve the system under H_B for $\tau/2$
5: Evolve the system under H_C for $\tau/2$

The order of the H_B and H_C steps is not important, as their Poisson bracket is equal to zero, so they commute. When performing many successive steps, this commutativity allows us to combine H_{Bi} and H_{Bi+1}, as well as H_{Ci} and H_{Ci+1}. For example, a short sequence of steps will look like

$$H_C(\tau/2)H_B(\tau/2)H_A(\tau)H_B(\tau)H_C(\tau)H_A(\tau)H_B(\tau/2)H_C(\tau/2) \tag{4.2.1}$$

Where $H_i(\tau)$ represents the evolution of the system under the i portion of the hamiltonian with timestep τ. This integrator is an implementation of the partitioned Euler method. In a more stripped down form, this method looks like

$$q_{n+1} = q_n + h\nabla_p H(p_n, q_{n+1}) \tag{4.2.2}$$
$$p_{n+1} = p_n + h\nabla_q H(p_n, q_{n+1}) \tag{4.2.3}$$

Which is simplified for the Hamiltonian splitting we have chosen, as for H_A and H_B we have

$$q_{n+1} = q_n + h\nabla_p H(p_n) \tag{4.2.4}$$
$$p_{n+1} = p_n + h\nabla_q H(q_{n+1}) \tag{4.2.5}$$

When compared to the standard Euler scheme 3.5.10, the only modification that was required to make our integrator symplectic was to change q_n for q_{n+1}. This dictates the order in which we evolve the co-ordinates under H_A. Figure 9 compares the trend in error of a Keplerian orbit when calculated using the simple (explicit) Euler and the symplectic partitioned Euler methods.

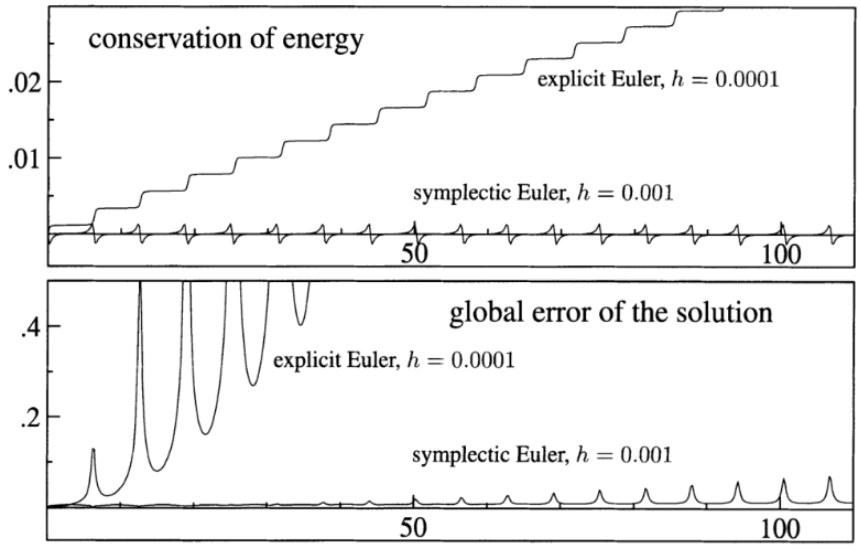

Figure 9: A comparison of the energy error associated with explicit and symplectic Euler integrators [11].

4.3 The Eliminator

The calculation of distances required to precisely determine whether or not two bodies are in the hybrid regime is computationally expensive. Since this must be done during the first half of the integration step, before the Keplerian step symplectically advances the positions of the bodies, determining exactly whether or not each body is in the hybrid regime each step will effectively double the compute time of the symplectic portion of the integrator. To mitigate this, an eliminator was employed to remove any bodies which we are absolutely sure will not be undergoing a close encounter during the next time step.

Since the eliminator will be employed at every step, it is important to choose an implementation which is computationally inexpensive. [7] suggests advancing the bodies on a Keplerian path. This would require performing two coordinate transformations every step, so it would have a noticeable impact on the compute time when integrating a small number of bodies. We instead choose a simple eliminator which advances the positions of each object linearly, checking to see if the estimated separations after this crude advancement are less than some predetermined value which assures none of the bodies are close enough to be in the hybrid regime, and passes the bodies which may be close enough to be in the hybrid regime to the more thorough distance checker in the H_B step. Crucially, this eliminator allows us to avoid making costly r_{ij} calculations for every $i \neq j$, which saves a significant amount of computations per step.

4.4 The Changeover Function

When switching between the symplectic and higher order regimes, we need some way to transfer the interaction terms from H_B over to H_A. For this purpose, we define a Hamiltonian scaling function $K(r_{ij})$. It will scale the different portions of the split Hamiltonian like

$$H_A = \sum_{i=1}^{N}(\frac{p_i^2}{2m_i} - \frac{Gm_\odot m_i}{r_\odot}) - \sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{m_i m_j}{r_{ij}}[1 - K(r_{ij})] \tag{4.4.1}$$

$$H_B = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\frac{m_i m_j}{r_{ij}}K(r_{ij}) \tag{4.4.2}$$

While H_C remains the same. When choosing the changeover function $K(r_{ij})$, we want to satisfy the following conditions [20]:

- When r_{ij} is large, $K = 1$, such that the interaction terms are handled by the symplectic integrator in H_B.
- When r_{ij} is small, $K = 0$, such that the interaction terms are handled completely by the higher order integrator along with the Keplerian terms in H_A.
- $K(r_{ij})$ is smooth.
- $K(r_{ij})$ is computationally inexpensive to evaluate.

A piecewise spline function was used. The function is defined as follows:

$$\begin{cases} K = 0 & y < 0 \\ K = \frac{y^2}{2y^2 - 2y + 1} & 0 < y < 1 \\ K = 1 & y > 1 \end{cases}$$

where

$$y = (\frac{r_{ij} - 0.1r_{crit}}{0.9r_{crit}}) \tag{4.4.3}$$

r_{crit} is a critical distance parameter which we define based on the system being integrated and the time step chosen, being some multiple of the mutual hill radii. In our integrations, it was chosen to be $3r_{mut}$. Duncan et Al [21] used this critical distance for their hybrid integrator when integrating the orbits of large planets. Since the planets in are dataset are estimated to be rather large, this seemed a good choice, though experimentation to find a better critical distance for our scale of problem to minimize the amount of time spent in the high order regime while also avoiding unnecessary contributions to error would be warranted. Figure 10 illustrates the shape of the function K.

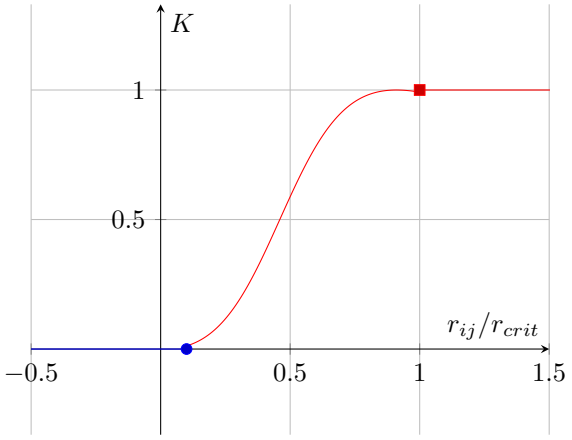

Figure 10: The spline function that smooths the transition from the hybrid to the symplectic regime.

4.5 The Higher Order Integrator

When the partitioned Euler integrator has to deal with close encounters in a Keplerian system, it loses the symplectic property. The energy error is no longer symmetric in the Keplerian q & p dimensions in phase space. It is for this reason that we must switch to a higher order integrator to maintain accuracy.

Our integrator of choice was Bulirsch-Stoer, implemented as described in Numerical Recipes in C [22]. This approach is based on the assumption that the solution returned by an integration step is some complicated function f of the stepsize τ. We perform the numerical calculation at various values of τ in order to probe the shape of $f(\tau)$, then fit it to some chosen analytic function, and evaluate the function at $\tau = 0$. Ideally, if we fit $f(\tau)$ well, this will be a good estimate of what the integrator would return if we had used an infinitesimally small stepsize.

Polynomial extrapolation was the method chosen for fitting $f(\tau)$ and evaluating at $\tau = 0$. The sequence of n's chosen, where $\tau = H/n$ defines the length of the substep size τ compared to the full step size H is

$$n = 2, 4, 6, 8, 10, 12, ..., [n_j = 2j], ... \tag{4.5.1}$$

Figure 11 illustrates the process of extrapolating $f(\tau)$ to $\tau = 0$. We successively reduce the size of the substeps, probing the shape of f, and then extrapolate to inf substeps.

Bulirsch-Stoer works well for smooth problems while being relatively computationally inexpensive. It does not work for problems where we expect singularities

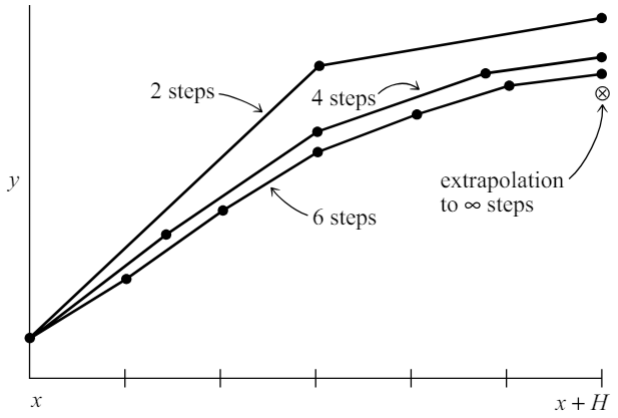

Figure 11: The richardson extrapolation based Bulirsch-Stoer method [22]

near or within the solution, or for stiff problems, where Runga-Kutta methods prove to be the better choice. This is fine four our problem, as we do not expect our planets to drift very close to each other; they should remain in their respective orbits. If we do detect a large error when performing a hybrid step (H_B term becomes unreasonably large) we can expect the orbital parameters to have changed significantly by the end of the integration, as this is indicative of a very close encounter taking place.

Algorithm 2 Hybrid Integrator Step

1: Use the eliminator to rule out any bodies without the potential to undergo a close encounter during this timestep
2: Evolve the system under H_C by shifting each body by $\tau \sum_i p_i / 2m_\odot$ for $\tau/2$
3: Evolve the system under H_B by;
- Determining which bodies are in a close encounter using radii from the last performed H_B step. Only calculate $K(r_{ij})$ exactly for bodies not eliminated, assume $K = 1$ otherwise.
- Giving each body the perturbative kicks of all other bodies but the central star for $\tau/2$, scaled by $K(r_{ij})$.

4: Evolve the system under H_A by;
- Giving all bodies not undergoing a close encounter a symplectic Keplerian kick for τ
- Advancing the position of each body not undergoing a close encounter, based on its momentum, for τ
- Placing all bodies undergoing a close encounter in the separate higher order integration scheme and evolving their H_A and $H_B(1 - K(r_{ij})$ terms independent of the bodies still in the Kepler regime
- Using calculated orbital radius measurement to find current hill radii for eliminator and H_B step

5: Evolve the system under H_B for $\tau/2$ by;
- Calculating the mutual distances between all bodies r_{ij}
- Giving each body the perturbative kicks of all other bodies but the central star for $\tau/2$, scaled by $K(r_{ij})$.

6: Evolve the system under H_C for $\tau/2$

There is a change in the way H_B is handled here, even during symplectic steps. Since the hybrid steps must begin at the end of a timestep, rather than partway through, we cannot directly combine every H_{Bi} with its neighbor H_{Bi+1}. Instead, since we know that H_C will not affect the *relative* position vectors involved in the interaction terms r_{ij} $\forall$ $i \neq j$, we can emit their calculation during the first H_B step, relying on the stored values from the previous step.

4.6 Fitting to TTVs

At the end of each integration, the goodness of fit of each planet's model transits needs to be evaluated against the real measured transits. χ^2 was used, where

$$\chi^2 = \sum_{i=1}^{n} \frac{(t_i - t_{0i})^2}{\sigma_i} \tag{4.6.1}$$

Here, t_i is the TTV of the i'th transit in the model (measured against the

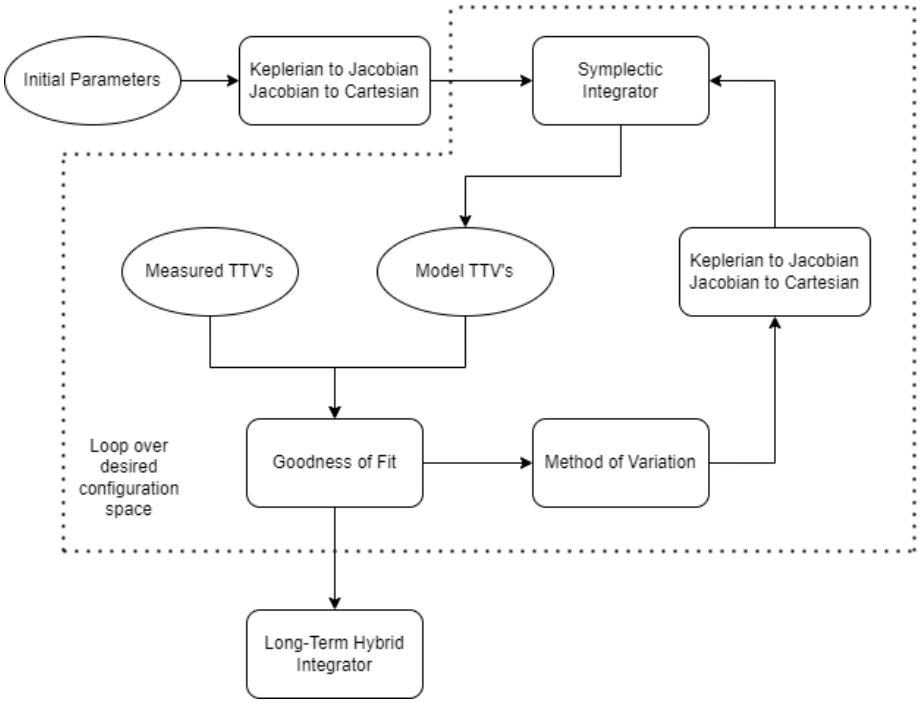

Figure 12: The entire analysis pipeline.

same expected value for the i'th real transit), and t_{0i} is the i'th transit TTV in the measured data. Since our integrator returns distinct points rather than a function for the positions of the objects, interpolation was required to determine the real transit time in our model, which was taken to be the time at which a body passed through the reference angle measured from the origin. In our integrator, this was simply $\hat{x}$. A small τ was used for this integrator, on the order of a few seconds, so linear interpolation was sufficient and did not introduce noise larger than a small fraction of a second to our model TTV data.

4.7 The Complete Scheme

Figure 12 illustrates how all of the modules of the complete scheme fit together. The general timeline is as follows.

1. An initial estimate of the Keplerian parameters of the planets in the system is inserted into the scheme.

2. The Kep-to-Jac & Jac-to-Cart modules translate the Keplerian co-ordinates into Cartesian for the integrator.

3. The integrator is run for the desired time or number of transits, while recording the transit times of each body.

4. The integrator returns the model TTVs, passing them to the goodness of fit module for comparison with the measured TTVs.

5. The initial parameters and goodness of fit are passed into the module to vary the initial parameters for the next integration.

6. The Kep-to-Jac & Jac-to-Cart modules translate the Keplerian parameters output by the variation module for the integrator.

This scheme is looped until the desired configuration space has been probed, or until a desired goodness of fit has been found, depending on the method of variation. We then run the long-term hybrid integrator on candidate configurations in order to rule out those which are unstable.

5 Data

The system tested by our scheme was K-385, known as KOI-2433 for the Kepler Space Telescope's mission. KOI-2433 is a sun-like star located in the constellation of Cygnus with a mass of $1.05m_{\odot}$, radius of $1.157R_{\odot}$, and temperature of $5829K$ [23]. It hosts seven close-orbiting transiting planets. This puts KOI-2433 up there with Trappist-1 in terms of the number of confirmed planets; and only one less than the record-holding Kepler-90 system. The planets have orbital periods ranging between a very short 3.37 and a Mercury-like 86.4 days. This means that constant dynamical interactions are present, which are ripe for exploration.

Since we are concerned with fitting TTVs, this is the data that was extracted from Kepler observations of our target system. The Kepler light curves were fitted, and the centre of transit time determined. The data was arranged into three columns; the expected transit time for a circular orbit C, the observed transit time O, and the error on 0. Data points for the transit times are provided for the majority of Kepler's mission [24]. Figure 13 is a sample plot for the TTVs and associated error bars for the closest planet, KOI-2433.8.
Missing transits were matched with their corresponding transit in the modelled data and thrown out by searching for jumps in the transit time larger than an orbital period plus some offset value to catch all of the positive TTVs.

Earth's approximate orbital data was taken from the Nasa Earth fact sheet [25] and used to validate the symplectic property of our integrator for an earth-like orbit.

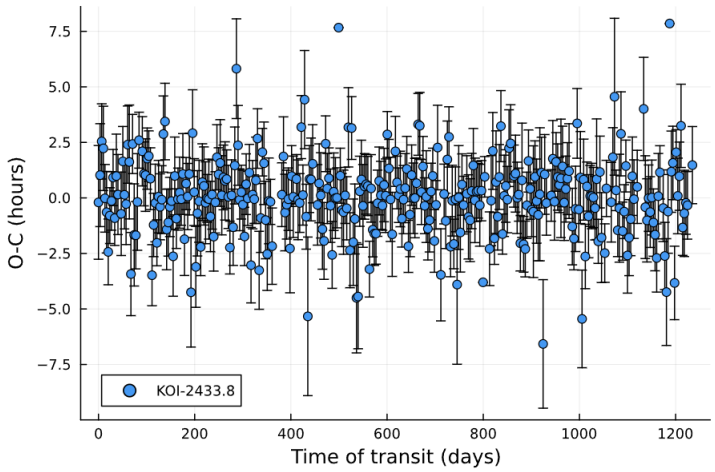

Figure 13: Observed-Calculated plot for the innermost planet KOI-2433.8

6 results

6.1 Validating our Integrator

In order to validate our kep-to-cart co-ordinate transformation module, a few sample orbits were plotted. Figure 14 is a comparison between the points returned by our integrator for the initial parameters $P = 1$, $e = 0.5$, $E = \pi/2$, and $\omega = \pi/2$, integrated for a single orbit with timestep $1 * 10^{-4}$ yr. Notice the starting position of the planet on the left hand side of the orbit as expected by the given E and ω.

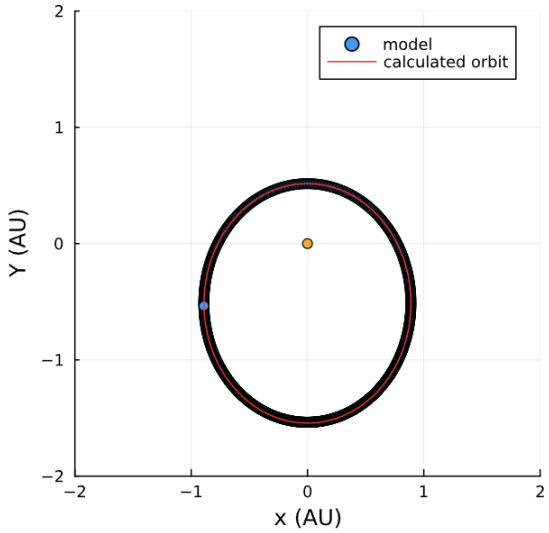

Figure 14: A sample orbit with initial conditions generated by the Kep-to-Cart module integrated by the symplectic integrator compared to the analytical solution.

This system was then integrated for a range of 2000 yr, with timestep $\tau = 2.5 * 10^{-}3$, in order to test the energy preserving qualities of our integrator. Figure 15 shows the fractional error of the Hamiltonian as the system evolves. This system is a particularly poor case for our class of integrator; it does not handle close encounters with the host star very well, which is the source of the spikes in our plot (In reality these spikes occur at much higher frequencies, one per orbit, and their frequency in this plot is an artifact of aliasing due to the resolution of our sampling). Still, we can see that the error is well bounded. There is no secular trend as our system evolves.

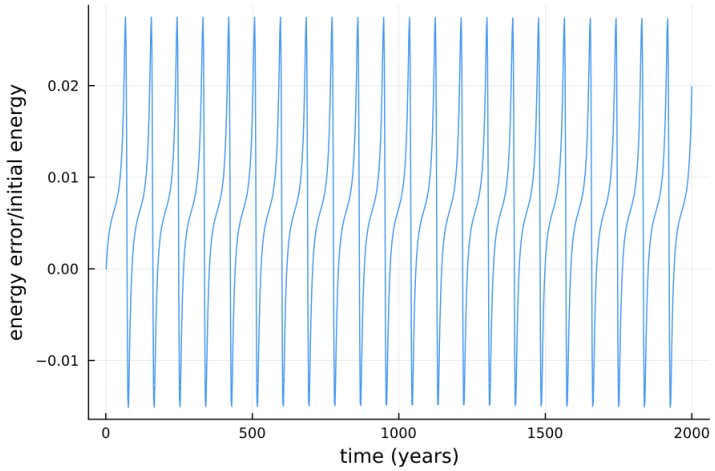

Figure 15: The proportional energy error of the sample orbit

For comparison, earth's orbit from J2000 [25] was calculated for the same time span using an identical time-step and the relative energy error plotted in figure 16. The total energy error is better controlled by a few orders of magnitude, and it remains bounded.

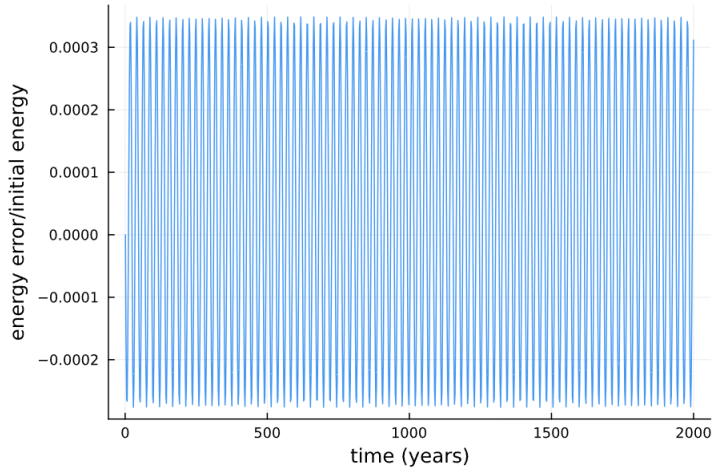

Figure 16: The proportional energy error of an earth-like orbit

Next, we prepared the target system K-385 with some rough parameter estimates. For the planets where mass estimates were available from the Exoplanet Archive [26], these estimates were used. For planets which were had no estimate, the mass was inferred from radius measurements by assuming an appropriate

density for the expected composition of a planet of a given size. The planets were then put in circular orbits with randomized phase and integrated for the length of the Kepler observations at a time step of one minute. The time taken to integrate these initial conditions was approximately 5 minutes.

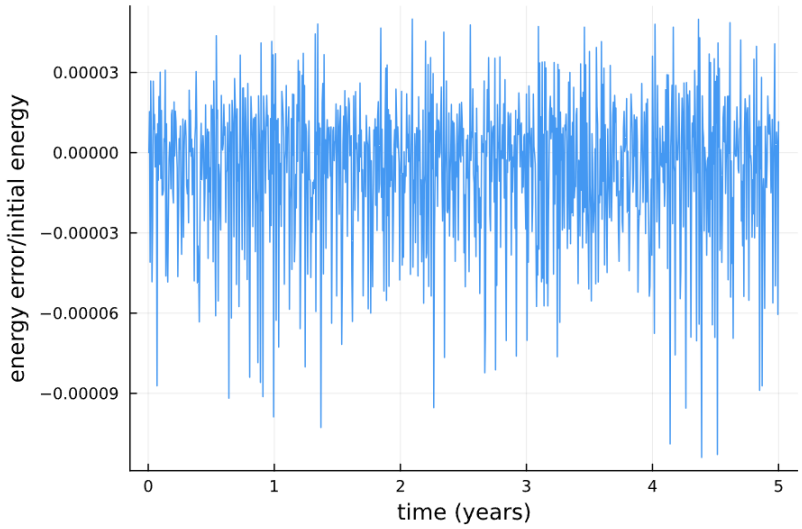

Figure 17: The proportional energy error of the KOI-2433 system when integrated over the full range of Kepler observations

The energy error is very well controlled, with $(E(t) - E_0)/E_0$ remaining on the order of $1 * 10^{-4}$ or lower when using a τ of $1min$ for the length of integration which we are concerned with. More important though, is the fact that there is no secular trend in the error over time. Longer integrations, out to 100 years, also preserved error bounds to the same order of magnitude.

6.2 Measurement-Model Fitting

Next, we begin to compare our model against the data for KOI-2433. During the integration, the times of transit with the reference angle are saved for each planet, and the real transit times are subtracted from the expected transit times of the Kepler transits. These are the Transit Timing Variations. These are plotted against the TTVs of the Kepler data. Using the same starting parameters as the last integration, plot 18 was generated for the fourth planet.

In order to see the effect of varying the initial parameters of the system on a planet's TTVs, we modified the phase of KOI-2433.2 which is the third planet in the system, while keeping the starting position of KOI-2433.1, which is the third planet in the system, the same. Planets 3 and 4 are in a 3:2 period resonance with each other, which exaggerates the cumulative effect of their interactions.

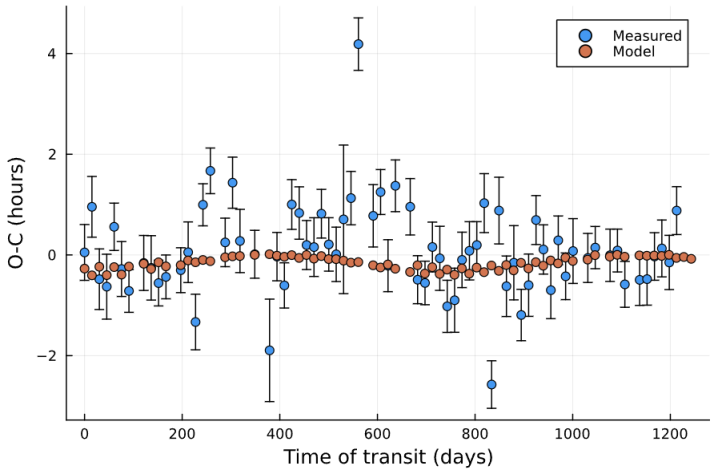

Figure 18: Planet 4 model vs measured TTVs with $E_4 - E_3 = 0$

Figure 19 is the planet 4 TTVs after changing planet 3's starting phase by π.

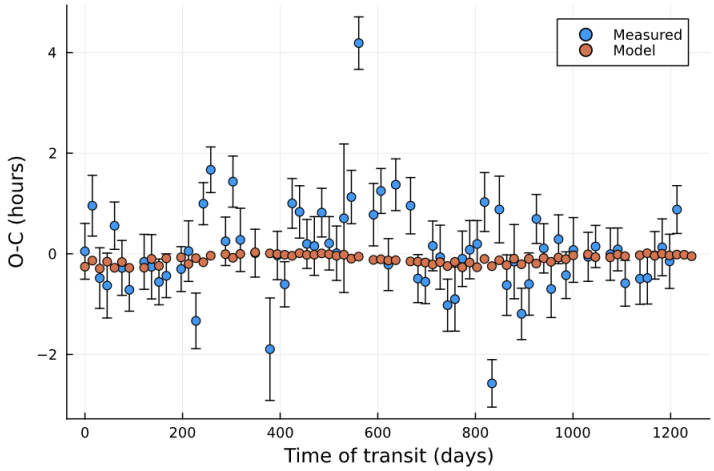

Figure 19: Planet 4 model vs measured TTVs with $E_4 - E_3 = \pi$

Figure 20 is the outcome of plotting the difference between these two TTV measurement plots. The beat in the TTV effect of 3 on 4 may indicate that we do not have the starting periods exactly right. Since it is always an approximation to assign a period to the orbit of an n-body system, it will take some trial and error to find the true resonant instantaneous period ratios. We verified that the TTVs visible in the plot for planet 4 are mostly due to the resonance with

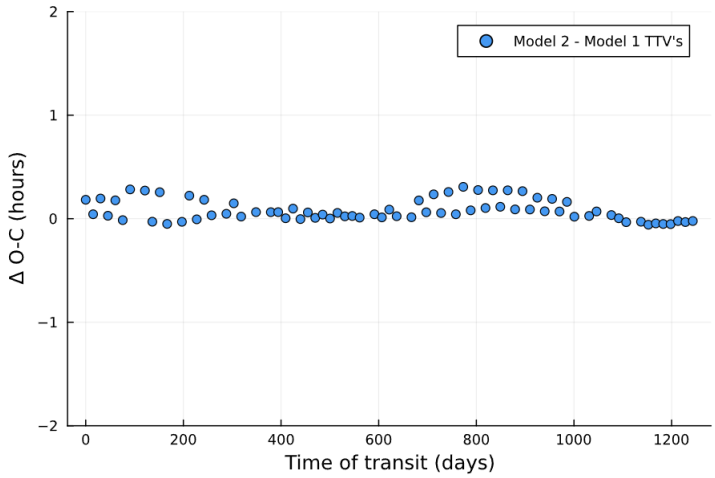

Figure 20: Plot of difference between $E_3 = 0$ and $E_3 = \pi$ models

planet 3 by setting planet 3's mass to zero. The resulting TTV plot was largely flat and showed no variations.

6.3 longterm integration

Figure 21 shows the relative energy error of a highly unstable system when integrated using the hybrid integration regime. The KOI-2433 planets were given unreasonably high e, randomized E, and the integrator was run for a period of 1000 years. The jumps in energy error correspond to close encounters that even our higher order integrator could not handle, likely collisions between the planets. This is the sort of behaviour we would expect to see in an unstable system. The orbital parameters returned at the end of the integration were wildly different from the starting parameters. Still, when there are no collisions, or integrator manages to maintain the symplectic property even in this highly dynamical system, and the plateaus between changes in energy are flat.

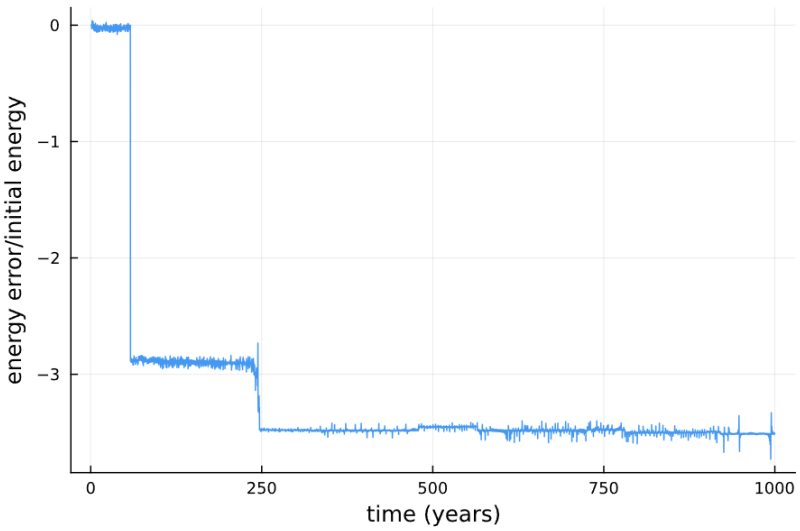

Figure 21: The behaviour of a highly unstable system

7 Conclusion

We set out to construct a hybrid symplectic integration scheme for finding the goodness of fit of system configurations to measured transit data. The methods used are largely a derivative of those developed by Wisdom and Holman [8], with some modifications, including the use of democratic heliocentric co-ordinates, as well as the addition of a higher order integrator to the symplectic scheme for running long-term stability analysis integrations with large time-steps.

The motivation behind the construction of our scheme was to provide a method to probe the configuration space of a target system in order to search for candidate configurations. With the use of statistical analysis routines, one could attempt to build likelihood distributions for the system's starting parameters by varying them and testing the goodness of fit of the resulting TTVs against real measured values.

The process of building this scheme first involved putting together a symplectic integrator capable of low-error orbit calculations. The method of Wisdom and Holman's [8] Modified Euler integrator, with a coordinate change to Democratic Heliocentric co-ordinates to speed up calculations, was employed. This integrator was built using Julia. It successfully manages to constrain the energy error of the integration to within 10^-4 for our target system when an appropriate time step is used. The number of calculations per time step is nonetheless kept low enough that integrations over the entire length of Kepler's observations can be done in a matter of minutes.

The next step was to build a method to compare the output of our model TTVs to the target system data. This involved matching model transits to their real counterparts, and throwing out any which were missing in the data. TTVs are compared to measurements using χ^2.

In order to allow the initial configuration parameters to be given in Keplerian co-ordinates, a co-ordinate conversion module was developed to transform between Keplerian and Cartesian co-ordinates, with Jacobi co-ordinates as an intermediate step. This module was used to generate sample systems which were compared to analytical Keplerian orbits. It successfully manages to initialize a range of orbits with different initial parameters which accurately follow the expected analytical paths.

Finally, a higher order integrator was added to the symplectic integrator to deal with close encounters when using larger timesteps. This addition allows us to use much larger timesteps when performing long term integrations. When close encounters occur, the bodies which are undergoing an encounter are handed over

to the higher order integrator in order to maintain accuracy. Bulirsch Stoer, implemented as described in section 4.5, was employed. The hybrid regime was then used to show the change in parameters of an unstable configuration over a long integration period.

While our scheme only deals with a reduced parameter space, our preliminary results show that we are still able to generate TTVs, especially when examining planets which are in resonant orbits. The next steps for this project involve the implementation of a method to vary the initial conditions of our model in order to begin to probe the parameter space of our problem.

8 Future Work

8.1 Statistical Methods

One of the initial goals of this thesis was to employ Bayesian analysis on the parameter space of our problem. While time constraints forced us to only perform a cursory test sampling, our pipeline is set up in such a way that a more thorough exploration of our target system's parameter space is possible. What we have constructed, with our entire scheme put together, is a method of sampling the error function associated with a given set of initial parameters, which can be fed into an MCMC algorithm in order to generate posterior distributions of our parameters. This would allow us to not only find the best fit parameters, but also an associated error.

Markov Chain Monte Carlo (MCMC) is a method of generating distributions which describe the likelihood of a certain parameter's value. These distributions are called the *posterior* in the parlance of Bayes theorem. To implement this algorithm, we would first use our error function χ^2 to create a likelihood function, for example, the Gaussian $e^{-\chi^2}$. This likelihood function is maximized when $\chi^2 = 0$. We then employ the Metropolis algorithm. This involves picking a point in parameter space $\vec{\theta}$, and calculating the likelihood associated with those parameters. We then pick another set of parameters θ'. The likelihood ratio between these points is $e^{-\chi^2_{\theta'}+\chi^2_{\theta}}$ which will be a real positive number between 0 and 1. A random number is picked in the same range, and compared to the likelihood ratio. If the ratio is larger, then we replace the current parameters with the new parameters and repeat the entire process. An instance of this process is called a *walker*. By initializing a series of these walkers, running them for a sufficient amount of time, and then examining their distribution in parameter space, we can build the posterior we were looking for. MCMC methods, paired with an n-body integrator to model dynamics, have been previously used to constrain the parameters of exoplanetary systems. [27] We hope to eventually be able to replicate the past success of this method using the integrator that was developed as described in this thesis.

8.2 Photometry

There is more information available in Kepler's data than was mentioned in this thesis. By taking full advantage of the light curve data, we could fit more than just transit times. Photometry, the fitting of the transit light curves, can be combined with dynamical modelling to better constrain model parameters. [27][28]. This method involves using the output of the N-body integrator, along with model parameters of the host star, to generate artificial light curves which can be analyzed and compared against measurements.

References

[1] A. Wolszczan and D. A. Frail, "A planetary system around the millisecond pulsar PSR1257 + 12", **355**, 145–147 (1992).

[2] A. Udalski, K. Zebrun, M. Szymanski, M. Kubiak, I. Soszynski, O. Szewczyk, L. Wyrzykowski, and G. Pietrzynski, "The Optical Gravitational Lensing Experiment. Search for Planetary and Low- Luminosity Object Transits in the Galactic Disk. Results of 2001 Campaign – Supplement", **52**, 115–128 (2002).

[3] Dec. 2015.

[4] C. D. Murray and S. F. Dermott, *Solar system dynamics* (Cambridge Univ. Press, 2010).

[5] G. M.-O. S.Seager, "A unique solution of planet and star parameters from an extrasolar planet transit light curve", The Astrophysical Journal **585**, `10.1086/346105` (2003).

[6] C. R. Nave, *Force on mass by spherical shell*, 2017.

[7] J. E. Chambers, "A hybrid symplectic integrator that permits close encounters between massive bodies", Monthly Notices of the Royal Astronomical Society **304**, 793–799 (1999).

[8] J. Wisdom and M. Holman, "Symplectic maps for the N-body problem.", **102**, 1528–1538 (1991).

[9] H. F. Levison and M. J. Duncan, "Symplectically Integrating Close Encounters with the Sun", **120**, 2117–2123 (2000).

[10] J. C. (cover), A. L. Garcia, B. K. (cover), A. Reeves, and P. F. Corey, *Numerical methods for physics (2nd edition)*, 2nd (Prentice-Hall, Inc., USA, 1999).

[11] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, Springer Series in Computational Mathematics (Springer Berlin Heidelberg, 2013).

[12] L. Landau and E. Lifshitz, *Mechanics: volume 1*, v. 1 (Elsevier Science, 1982).

[13] H. Yoshida, "Recent Progress in the Theory and Application of Symplectic Integrators", Celestial Mechanics and Dynamical Astronomy **56**, 27–43 (1993).

[14] D. Souami, J. Cresson, C. Biernacki, and F. Pierret, "On the local and global properties of gravitational spheres of influence", Monthly Notices of the Royal Astronomical Society **496**, 4287–4297 (2020).

[15] B. W. Carroll and D. A. Ostlie, *An Introduction to Modern Astrophysics*, edited by S. F. P. Addison-Wesley, 2nd (International) (2007).

[16] J. N. Winn, *Transits and occultations*, 2014.

[17] G. Wang and N. Espinoza, *A blind search for transit depth variability with tess*, 2023.

[18]A. C. Boley, C. Van Laerhoven, and A. P. Granados Contreras, "Transit duration variations in multiplanet systems", The Astronomical Journal **159**, 207 (2020).

[19]Almenara, J. M., Díaz, R. F., Bonfils, X., and Udry, S., "Absolute densities, masses, and radii of the wasp-47 system determined dynamically", AA **595**, L5 (2016).

[20]E. Agol, J. Steffen, R. Sari, and W. Clarkson, "On detecting terrestrial planets with timing of giant planet transits", **359**, 567–579 (2005).

[21]M. J. Duncan, H. F. Levison, and M. H. Lee, "A multiple time step symplectic algorithm for integrating close encounters", The Astronomical Journal **116**, 2067 (1998).

[22]W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in c*, Second (Cambridge University Press, Cambridge, USA, 1992).

[23]K. G. Stassun, R. J. Oelkers, M. Paegert, G. Torres, J. Pepper, N. De Lee, K. Collins, D. W. Latham, P. S. Muirhead, J. Chittidi, B. Rojas-Ayala, S. W. Fleming, M. E. Rose, P. Tenenbaum, E. B. Ting, S. R. Kane, T. Barclay, J. L. Bean, C. E. Brassuer, D. Charbonneau, J. Ge, J. J. Lissauer, A. W. Mann, B. McLean, S. Mullally, N. Narita, P. Plavchan, G. R. Ricker, D. Sasselov, S. Seager, S. Sharma, B. Shiao, A. Sozzetti, D. Stello, R. Vanderspek, G. Wallace, and J. N. Winn, "The Revised TESS Input Catalog and Candidate Target List", **158**, 138, 138 (2019).

[24]J. F. Rowe, S. T. Bryson, G. W. Marcy, J. J. Lissauer, D. Jontof-Hutter, F. Mullally, R. L. Gilliland, H. Issacson, E. Ford, S. B. Howell, W. J. Borucki, M. Haas, D. Huber, J. H. Steffen, S. E. Thompson, E. Quintana, T. Barclay, M. Still, J. Fortney, I. Gautier T. N., R. Hunter, D. A. Caldwell, D. R. Ciardi, E. Devore, W. Cochran, J. Jenkins, E. Agol, J. A. Carter, and J. Geary, "Validation of Kepler's Multiple Planet Candidates. III. Light Curve Analysis and Announcement of Hundreds of New Multi-planet Systems", **784**, 45, 45 (2014).

[25]D. R. Williams, *Earth fact sheet*, Jan. 2024.

[26]S. Hadden and Y. Lithwick, "Densities and Eccentricities of 139 Kepler Planets from Transit Time Variations", **787**, 80, 80 (2014).

[27]J. Sikora, J. Rowe, D. Jontof-Hutter, and J. J. Lissauer, "Refining the masses and radii of the star kepler-33 and its five transiting planets", The Astronomical Journal **164**, 242 (2022).

[28]D. Jontof-Hutter, J. Rowe, J. Lissauer, D. Fabrycky, and E. Ford, "The mass of the mars-sized exoplanet kepler-138 b from transit timing", English (US), Nature **522**, Funding Information: Acknowledgements D.J-H. acknowledges support through the NASA Postdoctoral Program and funding from the Center for Exoplanets and Habitable Worlds. The Center for Exoplanets and Habitable Worlds is supported by the Pennsylvania State University, the Eberly College of Science and the Pennsylvania Space Grant Consortium. J.F.R. acknowledges NASA grant NNX14AB92G issued through the

Kepler Participating Scientist Program. D.C.F. is an Alfred P. Sloane Fellow and was supported by the Kepler Participating Scientist Program award NNX14AB87G. E.B.F. was supported in part by NASA Kepler Participating Scientist Program award NNX14AN76G and NASA ExoplanetResearchProgramaward NNX15AE21G,aswellasthe Centerfor Exoplanets and Habitable Worlds. , 321–323 (2015).