American Journal of Computational Mathematics, 2013, 3, 195-204
http://dx.doi.org/10.4236/ajcm.2013.33028 Published Online September 2013 (http://www.scirp.org/journal/ajcm)

Jovian Problem: Performance of Some High-Order Numerical Integrators

Shafiq Ur Rehman[1,2]
[1]Department of Mathematics, The University of Auckland, Auckland, New Zealand,
[2]Department of Mathematics, University of Engineering and Technology, Lahore, Pakistan
Email: s.rehman@math.auckland.ac.nz, srehman@uet.edu.pk

Received July 3, 2013; revised August 1, 2013; accepted August 9, 2013

ABSTRACT

N-body simulations of the Sun, the planets, and small celestial bodies are frequently used to model the evolution of the Solar System. Large numbers of numerical integrators for performing such simulations have been developed and used; see, for example, [1,2]. The primary objective of this paper is to analyse and compare the efficiency and the error growth for different numerical integrators. Throughout the paper, the error growth is examined in terms of the global errors in the positions and velocities, and the relative errors in the energy and angular momentum of the system. We performed numerical experiments for the different integrators applied to the Jovian problem over a long interval of duration, as long as one million years, with the local error tolerance ranging from 10^{-16} to 10^{-8}.

Keywords: *N*-Body Simulations; Jovian Problem; Numerical Integrators; CPU-Time

1. Introduction

Computational astronomers make extensive use of accurate *N*-body simulations when studying the dynamics of the planets, asteroids and other small celestial bodies in the Solar System. These simulations are performed by first deriving a set of differential equations for the acceleration of the *N* bodies in the simulation, and specifying the initial positions and velocities of the bodies at time $t = t_0$. Generally, the initial value problems (IVPs) that occur for *N*-body simulations are a mixture of first- and second-order differential equations, but the sort of problems we are considering are of the form,

$$y''(t) = f\left(t, y(t)\right),\ y(t_0) = y_0,\ y'(t_0) = y_0', \qquad (1.1)$$

where $y_0 \in \mathbb{R}^k$ and $y_0' \in \mathbb{R}^k$ denote the initial positions and velocities, the operator denotes differentiation with respect to time t, and $f : \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}^k$ is a sufficiently smooth function. Here, k is the dimension of the IVP, which in some cases may change over time, as bodies are added or removed in the simulations. In some cases, these equations can be solved analytically, but mostly the differential equations are too complicated to find analytical solutions, necessitating the use of approximation techniques to find the numerical approximate solution. A wide range of integrators, for example, Runge-Kutta [3,4], Linear multistep [5], Runge-Kutta-Nyström [6], and Störmer [7] are used to find a numerical solution to the differential equations at $t = t_0 + ih$, with $i = 1, 2, \cdots$ and time-step h, which can depend on i.

2. Jovian Problem

The Jovian problem (see, for example, [1]) models the orbital motion of the Sun and the four Gas giants, Jupiter, Saturn, Uranus and Neptune, interacting through Newtonian gravitational forces. The Jovian problem is often used in numerical experiments, because the Gas giants collectively drive much of the dynamics of the Solar System. Let $r_i = [x_i, y_i, z_i]^T, i = 1, \cdots, 5$, be the position vector of the i^{th} body of the Jovian problem, where the bodies are ordered from Sun to Neptune and the coordinate system is the three-dimensional Cartesian system with the origin at the barycentre (centre of mass) of the bodies. Then the equations of motion for the i^{th} body can be written as

$$r_i''(t) = \sum_{j=1, j\neq i}^{5} \frac{\mu_j \left(r_j(t) - r_i(t)\right)}{\left\| r_j(t) - r_i(t) \right\|_2^3}, \quad i = 1, \cdots, 5, \qquad (1.2)$$

where $\|\cdot\|_2$ denotes the L_2-norm and μ_j is the gravitational constant G times the mass m_j of the j^{th} body, *i.e*, $\mu_j = Gm_j$. For each body we have a second-order differential equation for the x-, y-, and z-component,

giving us fifteen second-order differential equations in total. We express units of distance in astronomical units, the independent variable t in Earth days and the mass m_j in Solar mass.

We use the symmetry of interactions to reduce the number of calculations in the subroutine for evaluating the acceleration for the Jovian problem. Consider the individual terms in the summation and observe

$$\frac{\left(r_j(t)-r_i(t)\right)}{\left\|r_j(t)-r_i(t)\right\|_2^3}=\frac{-\left(r_i(t)-r_j(t)\right)}{\left\|r_j(t)-r_i(t)\right\|_2^3}.$$

Hence, once this term for r_j is calculated, we can update the acceleration for the second body by symmetry. Using this symmetry, we found that the subroutine for the evaluation of the force term for the Jovian problem reduces to approximately half of the CPU-time.

Unlike the Kepler problem, an analytical solution for the Jovian problem is unavailable. Therefore, numerical experiments using the Jovian problem require a reference solution in order to obtain an estimate of the error in the position and velocity. The reference solution has to be more accurate than the numerical solution. Since we plan to test the numerical integrators near the limit of double-precision arithmetic (2.2×10^{-16}), it is essential to use quadruple-precision arithmetic for the reference solution. Therefore, for long-term simulations, obtaining a reference solution can require considerable CPU-time.

Different types of errors are discussed throughout this paper. The global error is of major importance in the measurement of the quality of the numerical solution. We measure this global error in position and velocity, and also measure the relative error in energy and angular momentum. For the total error in the system the main source of error is the integration error, which consists of a truncation and round-off error. While performing accurate simulations, the round-off error contributes significantly to the global error because computers store numbers to only a certain precision. So, there will always be a loss of accuracy when performing long-term simulations. For fixed-step-size schemes, Brouwer [8] showed that, if the step-size is smaller than a prescribed value, the round-off error for conserved quantities, such as total energy and angular momentum, grows as $t^{\frac{1}{2}}$ and for other dynamical variables, such as coordinates of particles, as $t^{\frac{3}{2}}$. This error growth is known as Brouwer's law in the literature; see, for example, [9,10]. In contrast, when the round-off error is systematic, the power laws become t and t^2, respectively. In addition to these aspects, we investigate other effects of the round-off error here.

First we define the types of errors used in this paper. Let y_n and y_t be the vectors of the solution calculated numerically and the reference solution, respectively, and y_n' and y_t' are the vectors of the derivative to the numerical and reference solutions, respectively. Then the norm of the global errors in the position and the velocity are given by

$$E_r(t)=\left\|y_n-y_t\right\|_2,\ E_v(t)=\left\|y_n'-y_t'\right\|_2,$$

where, $\|\cdot\|_2$ is the unweighted L_2-norm.

Physical systems often have conserved quantities, for example, the total energy H or the total angular momentum L as for Kepler's two-body problem and the Jovian problem. Usually, these quantities will not be conserved exactly by the numerical solution and this derivation provides assessment about the accuracy of the solution. The total energy is defined as

$$H=\frac{1}{2}\sum_{i=1}^{N}m_i v_i^2-\sum_{j=1}^{N-1}\sum_{i=j+1}^{N}G\frac{m_i m_j}{d_{ij}},$$

where G is the gravitational constant, m_i the mass of the i^{th} body, v_i its velocity, and $d_{ij}=\| r_i-r_j \|_2$ the distance between the i^{th} and j^{th} bodies. The relative error in the energy can be calculated as

$$H_{rel}=\left|\frac{H_0-H}{H_0}\right|,$$

where H_0 is the total energy at the initial time $t=0$. However, we use

$$H_{rel}=\left|\frac{GH_0-GH}{GH_0}\right|,$$

to calculate H_{rel}, because the value of $\mu=Gm$ is known more accurately than m.

The total angular momentum L and the relative error of the angular momentum L_{rel} are defined as

$$L=\sum_{i=1}^{N}m_i r_i\times v_i,\ L_{rel}=\frac{\left\|L_0-L\right\|_2}{\left\|L_0\right\|_2},$$

where L_0 is the angular momentum at the initial time $t=0$. Note that a reference solution is not required to calculate H_{rel} and L_{rel}, unlike for the errors in the position and velocity. Hence, less computing resources are needed to measure the performance of the integrators here. However, H_{rel} and L_{rel}, being scalar quantities, impose only one constraint; if we look at the error in position or velocity then every component of E_r or E_v has to be small.

The phase error is the difference between the phase angle of the numerical solution and the reference solution. The phase error is defined by

$$\cos(\theta)=\frac{y_n\cdot y_t}{\left\|y_n\right\|_2\left\|y_t\right\|_2}.$$

where θ is the angle between the numerical solution and the reference solution.

3. Numerical Methods and Integrators

Explicit Runge-Kutta-Nyström methods (ERKN) were introduced by E. J. Nyström in 1925 [6]. The efficiency of an ERKN method depends upon the approach for controlling the error in the numerical approximations. One way of controlling the error is to use an adaptive step-size technique. In order to control the local error of a single step, a pair of formulae of different orders is used in such a way that the function evaluations of the two methods are identical. If the numerical solution y_n is obtained by using the lower-order formula, then the pair is said to be implemented in lower-order mode. However, it is recommended for efficiency reasons that the solution y_n be obtained using the higher-order formula for the next step [11], and the pair operated in this fashion is said to be implemented in higher-order mode or local extrapolation. In this paper we are using two variable-step-size ERKN integrators: Integrator *ERKN*689 is a nine stage, 6-8 FSAL pair [12] and integrator *ERKN*101217 is a seventeen stage, 10-12 non-FSAL pair [12].

3.1. Round-Off Error Control for ERKN Integrators

In this paper, we perform experiments with tolerance close to the machine precision (2.2×10^{-16}). Therefore, we investigate the possibility of reducing the growth of round-off error in the explicit Runge-Kutta-Nyström integrators using the technique known as compensated summation [13]. The idea of compensated summation is based on estimating the dominant contribution term of the round-off error. To explain the round-off error control technique, we consider the following solution formula

$$y_n = y_{n-1} + hy'_{n-1} + h^2\sum_{j=1}^{s} b_j K_j. \tag{1.3}$$

Equation (1.3) contains three types of errors; the integration error already in y_n from the previous time-step, the round-off error in the formation of hy'_{n-1} and $h^2\sum_{j=1}^{s} b_j K_j$, and the round-off error caused by adding terms on the right-hand side of (1.3). If the integration time-step is small then $hy'_{n-1} + h^2\sum_{j=1}^{s} b_j K_j$ is small compared to y_{n-1}. Hence, the round-off error will be dominated by adding $hy'_{n-1} + h^2\sum_{j=1}^{s} b_j K_j$ to y_{n-1}. In each time-step we estimate the round-off error caused by adding $hy'_{n-1} + h^2\sum_{j=1}^{s} b_j K_j$ to y_{n-1} and then update the solution as follows; First, calculate

$$\tau = hy'_{n-1} + h^2\sum_{j=1}^{s} b_j K_j - \delta,$$

where δ is the estimated round-off error on the previous time-step (at the start of the integration $\delta = 0$). Since $hy'_{n-1} + h^2\sum_{j=1}^{s} b_j K_j$ and δ are small compared to y_{n-1}, the error caused in the formation of τ is negligible. The solution is then updated to

$$Y_n = y_{n-1} + \tau,$$

and the estimated round-off error for the time-step is calculated as

$$\delta = Y_n - y_n - \tau \tag{1.4}$$

The solution is then updated as $y_n = Y_n$. The velocity formula also uses the same concept to control the round-off error.

We used the round-off error control technique to investigate the maximum error in position (E_r) and velocity (E_v) for the Jovian problem described in Section 2. The integration was performed over 10^6 years using $TOL = 10^{-2i}$, for $i = 4,5,\cdots,8$. **Table 1** shows the maximum values of E_r and E_v for the explicit Runge-Kutta-Nyström integrators *ERKN*689 and *ERKN*101217. The column labelled *With* contains E_r and E_v calculated when the integration is performed with round-off error control, whereas the column labeled *Without* contains the percentage variation corresponding to the values in column *With* when calculated E_r and E_v by performing integration with-out round-off error control.

For *ERKN*689, the maximum values for E_r and E_v with round-off error control are always less than E_r and E_v without round-off error control. The only exception is for $TOL = 10^{-10}$, where the values of E_r and E_v in the

Table 1. The maximum values of E_r and E_v for *ERKN*689 and *ERKN*101217 obtained with and with-out round-off error control applied to the Jovian problem over one million years for the local error tolerances 10^{-8}, 10^{-10}, 10^{-128}, 10^{-14}, 10^{-16}.

	***ERKN*689**				***ERKN*101217**			
	E_r		E_v		E_r		E_v	
TOL	*With*	*Without*	*With*	*Without*	*With*	*Without*	*With*	*Without*
10^{-8}	4.37×10^{-2}	+0.02%	6.31×10^{-5}	+0.02%	4.70×10^{-1}	+0.21%	6.78×10^{-4}	+0.29%
10^{-10}	1.11×10^{-4}	−0.91%	1.60×10^{-7}	−0.63%	1.63×10^{-3}	−0.62%	2.35×10^{-6}	−0.86%
10^{-12}	1.70×10^{-6}	+71.8%	1.94×10^{-9}	+68.9%	4.82×10^{-5}	−0.21%	6.63×10^{-8}	−0.15%
10^{-14}	9.74×10^{-7}	+86.0%	9.35×10^{-10}	+84.4%	2.42×10^{-5}	−6.61%	3.33×10^{-8}	−7.77%
10^{-16}	2.28×10^{-6}	+71.0%	1.58×10^{-9}	+78.5%	8.71×10^{-6}	−7.40%	1.19×10^{-8}	−6.25%

column *Without* are close to zero and insignificant compared with values for smaller tolerances. The maximum difference was observed for $TOL = 10^{-14}$. Here, the maximum values for E_r and E_v obtained with round-off error control were approximately 86% and 84% less than those obtained without round-off error control. For *ERKN*101217, except for $TOL = 10^{-8}$, we found that the round-off error control technique is not very effective, because the errors in the position and velocity obtained with round-off error control are not always less than E_r and E_v without round-off error control.

This could be because the average time-step for *ERKN*101217 over 10^6 years is quite large. For example, with $TOL = 10^{-14}$, *ERKN*101217 takes a time-step of approximately 144 days on average over 10^6 years, and hence, the assumption that $hy'_{n-1} + h^2 \sum_{j=1}^{s} b_j K_j$ is small relative to y_{n-1} is invalid. Therefore, for *ERKN*101217, using $TOL = 10^{-2i}$, with $i = 5,6,7,8$, it is not recommended to use the round-off error control technique.

3.2. *ODEX*2 Integrator

For the direct numerical solution of systems of second-order ordinary differential equations, Hairer, Nørsett and Wanner [14] developed an extrapolation code *ODEX*2 based upon the explicit midpoint rule with order selection and step-size control. The *ODEX*2 integrator is good for all tolerances, especially for high precision, like 10^{-20} or 10^{-30}. To observe the change in results for E_r, we performed experiments with a variety of default settings of *ODEX*2, for example, by setting the parameter *ITOL* used for controlling the local error to 0 or 1. We observed that there is hardly any significant difference in results when applied to the Jovian problem over one million years for $TOL = 10^{-16}$ to 10^{-8}.

3.3. Step-Size Variation

Here, we investigate the step-size variation for the variable-step-size integrators *ERKN*689, *ERKN*101217, and *ODEX*2 applied to the Jovian problem. The eccentricities of the orbits of the Jovian planets are no more than 0.1 and there are no close-encounters between the planets. Therefore, the variable-step-size integrators should require small step-size variation. **Table 2** shows the step-size variation for the above integrators applied to the Jovian problem over one million years for the local error tolerances in the range 10^{-16} to 10^{-8}. The columns h_{mn} and h_{mx} list the percentage variation in the minimum and maximum step-sizes relative to the mean step-size. For example, h_{mn} is calculated as

$$h_{mn} = 100\left(\frac{h_{min} - \bar{h}}{\bar{h}}\right),$$

where, h_{min} is the smallest step-size used and $\bar{h}$ the mean step-size. For these results, we considered the on-scale step-sizes by ignoring the first few step-sizes in a transient region near $t = 0$ as well as the final step-size.

The step-size variation depends both upon the integrator and the tolerance chosen and ranges from approximately −34% to 152%. The largest variation between the maximum and minimum step-sizes occurs for *ERKN*689 with $TOL = 10^{-16}$, where it is a factor of three, with h ranging from $0.89\bar{h}$ to $2.52\bar{h}$. For the purpose of our work, we regard this variation as small. This small step-size variation enables us to add a fixed-step-size integrator $\bar{S}$ -13 (Störmer of order 13) in this paper. Therefore, we conclude that *TOL* has little effect on the step-size variation for *ERKN*101217 and *ODEX*2.

To see the effect of round-off error, we also performed integrations with $TOL = 10^{-14}$ in quadruple-precision arithmetic. The percentage variations of h_{mn} and h_{mx} were approximately −18% and 133% for *ERKN*689, −20% and 21% for *ERKN*101217, and −30% and 21% for *ODEX*2. Except for h_{mx} in *ERKN*101217, the step-size variations obtained in quadruple-precision arithmetic have reasonably good agreement with **Table 2**. Hence the round-off error is not significant with $TOL = 10^{-14}$.

3.4. Störmer Methods

Störmer methods are an important class of methods for solving systems of second-order differential equations. Introduced by Störmer [7], the methods have long been utilised for accurate long-term simulations of the Solar System [2]. Grazier [15] recommended an order-13, fixed-step-size Störmer method that uses backward differences

Table 2. Step-size variation for the variable-step-size integrators *ERKN*689, *ERKN*101217, and *ODEX*2 applied to the Jovian problem over one million years, with the local error tolerance *TOL* as specified in the first column.

	***ERKN*689**		***ERKN*101217**		***ODEX*2**	
TOL	h_{mn}	h_{mx}	h_{mn}	h_{mx}	h_{mn}	h_{mx}
10^{-8}	−17%	84%	−20%	23%	−30%	14%
10^{-10}	−17%	99%	−20%	22%	−13%	12%
10^{-12}	−18%	115%	−19%	21%	−30%	31%
10^{-14}	−18%	134%	−20%	32%	−29%	21%
10^{-16}	−18%	152%	−34%	71%	−21%	26%

in summed form, summing from the highest to lowest differences. The test results in [9] for simulations of the Sun, Jupiter, Saturn, Uranus and Neptune in double precision showed that the error in the energy and phase error grows as $t^{1/2}$ and $t^{3/2}$, respectively, to within numerical uncertainty when the step-size is $(\frac{1}{1024})$-th (4.1 days) of Jupiter's orbital period. This choice of step-size ensures that the local truncation error of the Störmer method is well below machine precision. In this paper we consider the fixed-step-size Störmer method of order 13 and refer to it as the $\bar{S}$-13 integrator.

4. Numerical Experiments for Long-Term Simulation

First we consider the error growth in the position and velocity using the variable-step-size integrators *ODEX*2, *ERKN*689, and *ERKN*101217. We obtained the reference solution in quadruple-precision using *ERKN*101217 with $TOL = 10^{-18}$. To justify this particular choice for the reference solution, we integrated the Jovian problem using the *ERKN*101217 integrator with $TOL = 10^{-20}$. The maximum difference between the positions and velocities of these two solutions is no more than 4.61×10^{-13}. We also integrated the Jovian problem in quadruple-precision with the tolerance $TOL = 10^{-18}$, but using the *ERKN*689 integrator and found that the maximum difference with the solution for *ERKN*101217 with $TOL = 10^{-18}$ is no more than 5.11×10^{-13}. This suggests that the *ERKN*101217 integrator with $TOL = 10^{-18}$ is sufficiently accurate to obtain the reference solution.

Figure 1 illustrates the unweighted L_2-norm of the estimation of the maximum global error in the position as a function of tolerance with three variable-step-size integrators *ODEX*2, *ERKN*689, and *ERKN*101217 for the Jovian problem over one million years. In most cases, the maximum global error occurs at the end point of the integration. We chose the range 10^{-16} to 10^{-8} for the local error tolerance, because 10^{-16} is close to machine precision in double-precision arithmetic and tolerances greater than 10^{-8} lead to errors that are too large to be meaningful. The result for *ODEX*2 is an accuracy (maximum global error) that ranges from 7.4×10^{-5} to 1.1×10^{2}. We observe that the maximum accuracy (minimum of the maximum global error) is obtained with $TOL = 10^{-16}$ and the minimum accuracy with $TOL = 10^{-8}$. The graph for *ODEX*2 exhibits three phases: In the middle phase, with $TOL \in [10^{-15}, 10^{-12}]$ the round-off error does not yet affect the global error. Round-off has an effect for $TOL \le 10^{-16}$, which we further investigated by using smaller tolerances, *i.e*, $TOL = 10^{-16+k}$, for k = 0.2, 0.4, 0.6, 0.8, and 1. As TOL decreases further from 10^{-16}, the global error starts to increase again, which indicates the influence of the round-off error. The phase for $TOL > 10^{-11}$, shows a global error of approximately 10^{1} AU, which is the diameter of Jupiter's orbit. Here, the integrator still finds the orbit but at an arbitrary position angle that could deviate as much as 180°. We evaluated the phase error using the formula described in Section 2 and found that it is approximately 172°. This means that the amplitude of the orbit is not changing, but the error in its phase angle may be as large as π.

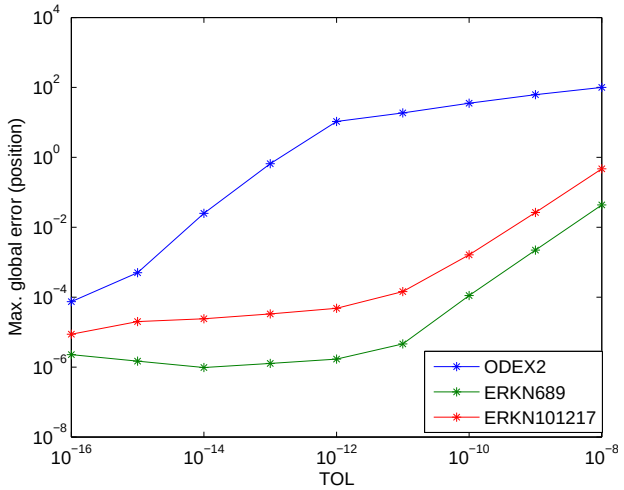

Figure 1. The maximum global error in position for the variable-step-size integrators *ODEX*2, *ERKN*689, and *ERKN*101217 applied to the Jovian problem over one million years for local error tolerances ranging from 10^{-16} to 10^{-8}.

Let us now consider the integrator *ERKN*101217 in **Figure 1**. Here, the accuracy ranges from 8.7×10^{-6} to 4.6×10^{-1}. The maximum accuracy is again obtained with $TOL = 10^{-16}$ and the minimum with $TOL = 10^{-8}$. We observe that from $TOL = 10^{-11}$ to 10^{-16} there is hardly any gain in accuracy. Therefore, if the best accuracy is required then $TOL = 10^{-16}$ should be used, but otherwise, a small sacrifice in accuracy will save a considerable amount of CPU-time.

The integrator *ERKN*689 has an accuracy ranging from 9.7×10^{-7} to 4.4×10^{-2}, with the maximum accuracy obtained at $TOL = 10^{-14}$ and the minimum at $TOL = 10^{-8}$. Therefore, nothing is gained by decreasing the tolerance from 10^{-14} to 10^{-16}. The maximum at $TOL = 10^{-14}$ is an indicator that the round-off error affects the global error when using tolerances between 10^{-14} and 10^{-16}. To measure the possible effect of round-off error, we performed experiments in quadruple-precision. We obtained the maximum global error in the position as a function of tolerance for the local error tolerances 10^{-16} and 10^{-10} using *ERKN*689 and *ERKN*101217. We observed that both curves are straight and maintain a difference of about 1.5 orders of magnitude. In particular, the graph is not bending up for *ERKN*689 using the small tolerance of 10^{-16}. This confirms the effect of round-off

error in the double-precision arithmetic. We conclude from **Figure 1** that, for local error tolerances ranging from $TOL = 10^{-16}$ to 10^{-8}, the integrator *ERKN*689 is the most accurate and *ODEX*2 is the least accurate integrator.

Let us now compare this performance with the $\bar{S}$-13 integrator. **Figure 2** shows the error growth in the position for the Jovian problem using the integrators $\bar{S}$-13, *ODEX*2, *ERKN*689, and *ERKN*101217. The integration was performed over 10^6 years and the error was sampled at every 100 years. The integration with the $\bar{S}$-13 integrator was performed in double-precision using a step-size of four days.

We performed two sets of experiments. For the first set of experiments, we maintained a given accuracy of approximately 10^{-4} for the maximum global error in the position over 10^6 years. We set $TOL = 10^{-16}$, 10^{-10}, and 10^{-11} for *ODEX*2, *ERKN*689, and *ERKN*101217, respectively; note that this variation in tolerance is necessary to achieve the prescribed accuracy, as illustrated in **Figure 1**. For small t, *ERKN*689 and *ERKN*101217 are more accurate than the other two integrators, but there is a crossover approximately at 5×10^4 years. We see in **Figure 2** that the three variable-step-size integrators achieve almost the same accuracy for the global error in position at the end of 10^6 years of integration and the fixed-step-size integrator $\bar{S}$-13 achieves almost one order of magnitude better accuracy than the variable-step-size integrators.

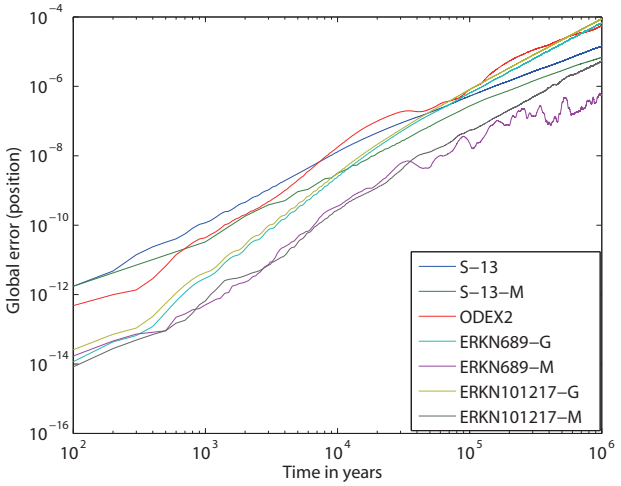

Figure 2. The error growth in position for the integrators $\bar{S}$-13, *ODEX*2, *ERKN*689, and *ERKN*101217 applied to the Jovian problem over one million years. The selection of local error tolerances is subject to a prescribed maximum accuracy.

To gain insight about the error growth depicted in **Figure 2**, we used unweighted linear least squares to fit a power law αt^{β} to E_r. We found that the integration error for the integrators *ERKN*689 and *ERKN*101217 grows approximately as t^2 (quadratic growth), while for *ODEX*2 and $\bar{S}$-13 it is approximately $t^{3/2}$. The error growth for *ODEX*2 is unexpected. Therefore, we repeated the integrations for *ODEX*2 by increasing the tolerance from $TOL = 10^{-16}$ to 10^{-15} and 10^{-14}; then we observe approximately the quadratic error growth.

The second set of experiments for integrators $\bar{S}$-13, *ERKN*689, and *ERKN*101217, labelled *S*-13-M, ERKN689-M and ERKN101217-M in **Figure 2**, respectively, are done such that maximum accuracy is maintained. To attain maximum accuracy, integrators *ERKN*689 and *ERKN*101217 use $TOL = 10^{-14}$ and 10^{-16}, respectively. For $\bar{S}$-13, we performed experiments with step-size variations as shown in **Table 3**. We observe that $\bar{S}$-13 achieves best accuracy with a step-size of approximately 10 days. The performance of the *ODEX*2 integrator at the prescribed accuracy, as shown in **Figure 1**, is also at the maximum accuracy for the local error tolerance of 10^{-16}. When performed at the maximum accuracy, there is no longer a crossover of the $\bar{S}$-13 integrator with the integrators *ERKN*689 and *ERKN*101217. At the end of 10^6 yuracyears of integration, *ERKN*689 achieves the best acc and *ERKN*101217 achieves the next best accuracy.

Some of the plots in these kinds of experiments have high-frequency oscillations. In order to smooth that data, the filter command in Matlab was employed with a window size of 50. The appropriate choice of window size is important. We have experimented (using the experiments illustrated in **Figure 2** with the exclusion of those labelled S-13M, ERKN689-M, and ERKN101217-M) for values of window sizes, 0, 10, 20, and 50 as shown in **Figure 3**. **Figure 3(a)** shows the result without filtering (WS = 0). There are enough oscillations of sufficient amplitude that it is difficult to distinguish the graphs. If the window size is small, as shown in **Figure 3(b)** (WS = 10) then quite a few oscillations are still there and it is not clear which of the integrators is being crossed. A window size of 50 seems to be a sensible value for this set of experiments. As is shown in **Figure 3(d)**, it is quite clear that $\bar{S}$-13 crosses only the integrators *ERKN*689 and *ERKN*101217. We also observed (although not shown)

Table 3. The maximum global error as a function of the step-size for the fixed-step-size integrator $\bar{S}$-13, applied to the Jovian problem over one million years.

Days	4	10	15	20	25	30
Global error in position	1.96×10^{-5}	1.08×10^{-5}	1.89×10^{-5}	3.95×10^{-5}	6.23×10^{-5}	1.05×10^{-4}

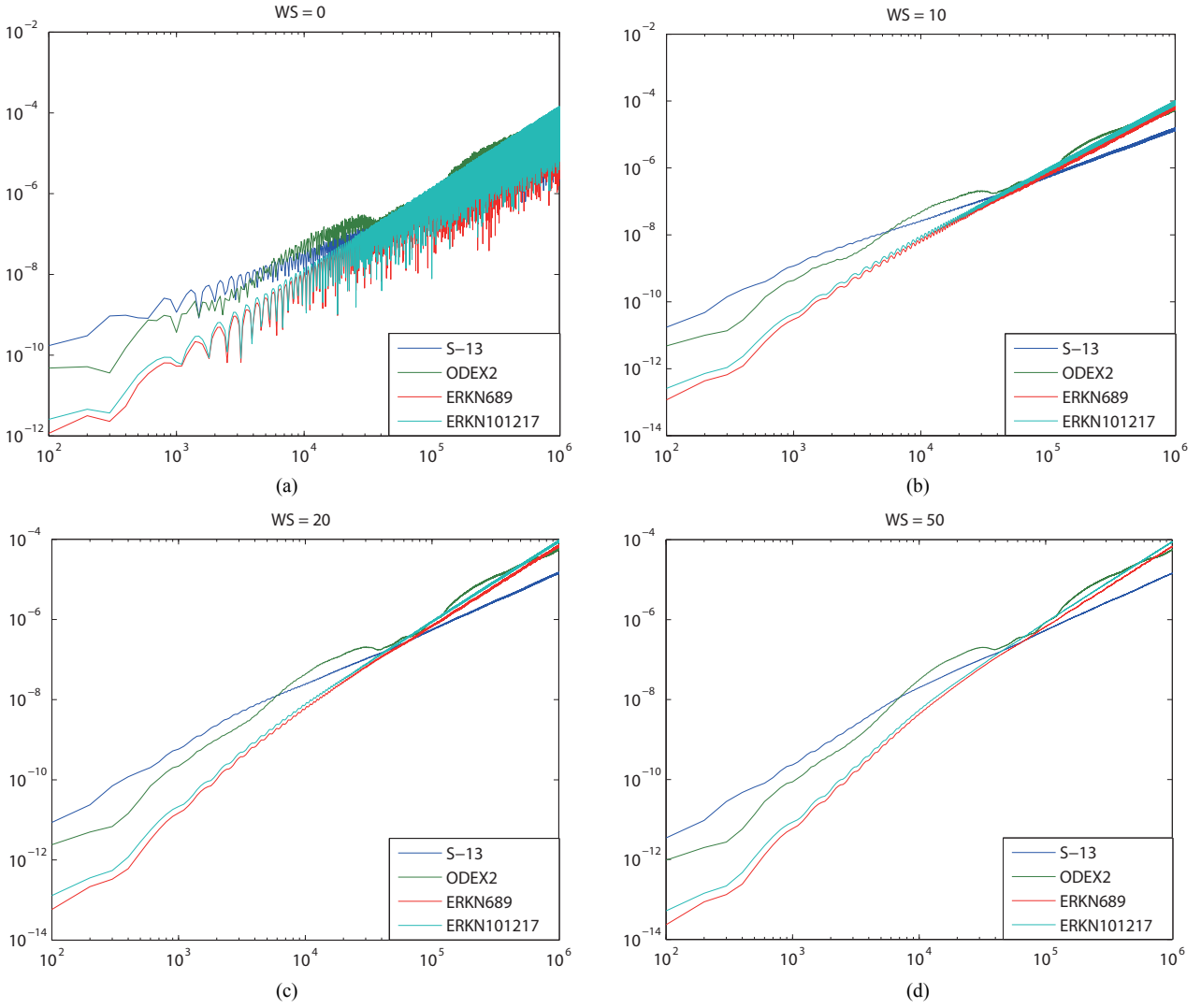

Figure 3. Experiments with a variation of the window size for the Matlab function filter. The window size is set to 0, 10, 20 and 50 in plots (a)-(d), respectively.

that filtering can complicate the interpretation of results for the first WS points, but this effect can be removed by ignoring the first WS points.

Let us now consider the accuracy of the integrators in terms of the relative error in energy and angular momentum. **Figure 4** shows the error growth in the energy for the Jovian problem. The integration has been performed in double-precision over 10^6 years using the same local error tolerances and integrators as for the results shown in **Figure 2**. For this set of experiments, we used the filter command in Matlab with a larger window size of WS = 100, because the oscillations were more pronounced than the set of experiments shown in **Figure 2**. The interval of integration is divided into 10,000 evenly spaced sub-intervals. To see the effect on the performance of the integrator by forcing it to hit every 100 years. We also performed experiments using *ERKN*101217, where we forced the integrator to hit every 50 and 200 years. We found three parallel graphs with a maximum difference in errors at 10^6 years of no more than 3.5×10^{-13}. Using 10,000 sub-intervals, we calculate the L_2-norm of the relative error in energy and angular momentum on the last accepted time step at the end of each sub-interval.

Similar to the set of experiments illustrated in **Figure 2** that attain a given accuracy of 10^{-4}, for the integrators *ERKN*689 and *ERKN*101217 (labeled by ERKN689-G and ERKN101217-G in **Figure 4**, respectively) we observe an error growth proportional to t in energy and angular momentum. For *ODEX*2, the error growth for energy and angular momentum shows some oscillations. The integrations were repeated for *ODEX*2 by increasing the tolerance from $TOL = 10^{-16}$ to 10^{-15} and 10^{-14}, which causes the oscillations to disappear. This indicates that round-off error is the cause of the oscillations. Approximately linear error growth in energy and angular

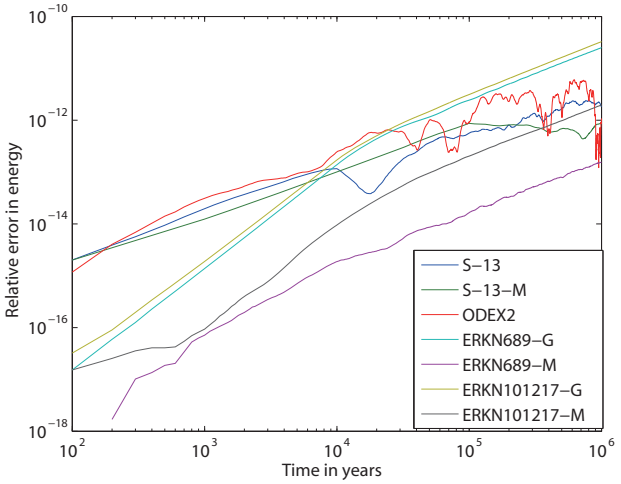

Figure 4. The error growth in the energy for the four integrators $\bar{S}$ -13, *ODEX*2, *ERKN*689, and *ERKN*101217 applied to the Jovian problem over one million years. The selection of local error tolerances is subject to attaining the given and maximum accuracy.

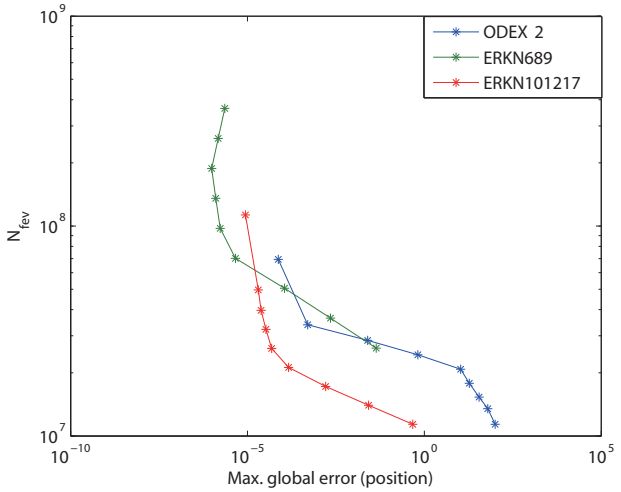

Figure 5. Efficiency plots showing the number N_{fev} of function evaluations against the L_2-norm of the maximum global error in position, obtained for the variable-step-size integrators *ERKN*689, *ERKN*101217, and *ODEX*2, applied to the Jovian problem over one million years with *TOL* ranging from 10^{-16} to 10^{-8}.

momentum was observed particularly for *ODEX*2 with $TOL = 10^{-14}$. As in **Figure 2**, the integrators *ODEX*2 and $\bar{S}$ -13 with step-sizes of four days, cross the integrators *ERKN*689 and *ERKN*101217.

However, this crossover for the relative error in energy occurs at a smaller t than for the global error in position. We observe from **Figure 4** that, for the relative error in energy, the integrator ERKN689 using $TOL = 10^{-14}$ (labeled by ERKN689-M) again achieves the best accuracy.

Let us now consider the efficiency of the integrators, which is the amount of work to attain prescribed accuracy. One way of measuring the work of different integrators is to count the number of function evaluations. **Figure 5** shows plots of the number of function evaluations against the maximum global error in position, obtained for the variable-step-size integrators *ERKN*689, *ERKN*101217, and *ODEX*2, and applied to the Jovian problem over one million years with TOL ranging from 10^{-16} to 10^{-10}. As described in **Figure 1** the best accuracy for ERKN689 is achieved at $TOL = 10^{-14}$, which needs approximately 1.7 and 2.7 times more function evaluations than *ERKN*101217 and *ODEX*2, respectively. If we consider tolerances such that all three integrators achieve the same accuracy 10^{-4} then *ERKN*101217 is the most efficient, because it uses the least number of function evaluations. The integrator *ERKN*689 is approximately 2.4 and *ODEX*2 approximately 3.3 times more expensive than *ERKN*101217. Our conclusion slightly changes by reducing the accuracy from 10^{-4} to approximately 10^{-3} or 10^{-2}. The integrator *ERKN*101217 again achieves the best accuracy compared to the integrators *ODEX*2 and *ERKN*689. For an accuracy of approximately 10^{-3}, the integrator *ODEX*2 is approximately 1.9 and *ERKN*689 approximately 2.1 times more expensive than *ERKN*101217. In contrast, for an accuracy of approximately 10^{-2}, the integrators *ODEX*2 and *ERKN*689 achieve almost the same accuracy and are approximately 2 times more expensive than *ERKN*101217.

We also investigated the CPU-time taken by the same variable-step-size integrators applied to the Jovian problem over one million years with local error tolerances in the range from 10^{-16} to 10^{-8}. For $TOL = 10^{-10}$, we found that *ODEX*2 and *ERKN*101217 take almost the same CPU-time, but *ERKN*101217 has approximately four orders of magnitude better accuracy than *ODEX*2. For the same tolerance, *ERKN*689 is almost three times more expensive than *ERKN*101217 and *ODEX*2, but has approximately one and five orders of magnitude better accuracy, respectively. For a given accuracy of approximately 10^{-4}, 10^{-3}, and 10^{-2}, *ERKN*101217 takes the least CPU-time. Hence, the integrator *ERKN*101217 is the cheapest option.

For the given range of tolerances from 10^{-16} to 10^{-10}, we found that *ERKN*689 achieves the best accuracy (at $TOL = 10^{-14}$), which is approximately one and two orders of magnitude better than the best accuracies achieved by *ERKN*101217 and *ODEX*2, respectively. At the same point in-time, *ERKN*689 is almost 1.6 and 2.4 times more expensive than *ERKN*101217 and *ODEX*2, respectively. These results clearly illustrate a trade-off between accuracy and efficiency.

5. Conclusions

The main objective of this paper was to analyse and

compare the efficiency and the error growth for different numerical integrators applied to the realistic problem involving the Sun and four Gas-giants. Throughout the paper, we examined the growth of the global error in the positions and velocities of the bodies, and the relative error in the energy and angular momentum of the system. The simulations were performed over as much as 10^6 years.

For long-term simulations, we performed experiments to observe the error growth in the positions and velocities using the variable-step-size integrators *ODEX*2, *ERKN*689, and *ERKN*101217, applied to the Jovian problem over one million years for local error tolerances in the range 10^{-16} to 10^{-8}. We observed that the integrators *ODEX*2, *ERKN*689, and *ERKN*101217 attained maximum accuracy with $TOL = 10^{-16}, 10^{-14}$, and 10^{-16}, respectively. Overall, we observed that for the local error tolerances in the range $TOL = 10^{-16}$ to 10^{-8}, the integrator *ERKN*689 is the most accurate and *ODEX*2 is the least accurate. We also observed that the integration error for the integrators *ERKN*689 and *ERKN*101217 grows approximately as t^2, while it grows as $t^{3/2}$ for *ODEX*2 and $\overline{S}$ -13. The error growth for *ODEX*2 was unexpected. Therefore, integrations were repeated for *ODEX*2 by increasing the tolerance from $TOL = 10^{-16}$ to 10^{-15} and 10^{-14}, for which we did observe the quadratic error growth.

We then investigated the efficiency of the integrators by counting the number of function evaluations against the maximum global error. We observed that the best accuracy achieved by *ERKN*689 uses approximately 1.7 and 2.7 times more function evaluations than *ERKN*101217 and *ODEX*2, respectively. Instead, if we require approximately the same accuracy of 10^{-4} achieved by all three integrators, the *ERKN*101217 is the most efficient, because it uses the least number of function evaluations. The integrator *ERKN*689 is approximately 2.4 and *ODEX*2 approximately 3.3 times more expensive than *ERKN*101217. We then investigated the CPU-time and observed that for a given accuracy of 10^{-4}, the number of function evaluations is proportional to the CPU-time. Hence, also in terms of CPU-time *ERKN*101217 is the cheapest option, which is approximately 2.4 and 3.3 times more efficient than *ERKN*689 and *ODEX*2, respectively. For the given range of tolerances from 10^{-16} to 10^{-8}, the integrator *ERKN*689 achieved best accuracy, which is approximately one and two orders of magnitude better than the best accuracy achieved by *ERKN*101217 and *ODEX*2, respectively. At the same point in time, *ERKN*689 is almost 1.6 and 2.4 times more expensive than *ERKN*101217 and *ODEX*2, respectively. These results clearly illustrate a trade-off between the accuracy and the efficiency.

We also measured the accuracy of the integrators by obtaining the relative error in energy and angular momentum. For the integrators *ERKN*689 and *ERKN*101217, the error growth is proportional to t, and for *ODEX*2 with $TOL = 10^{-14}$, we observe approximately linear error growth in energy and angular momentum.

6. Acknowledgements

The author is grateful to the Higher Education Commission (HEC) of Pakistan for providing the funding to carry out this research. Special thanks go to Dr. P. W. Sharp and Prof. H. M. Osinga for their valuable suggestions, discussions, and guidance throughout this research.

REFERENCES

[1] P. W. Sharp, "*N*-Body Simulations: The Performance of Some Integrators," *ACM Transactions on Mathematical Software*, Vol. 32, No. 3, 2006, pp. 375-395. doi:10.1145/1163641.1163642

[2] K. R. Grazier, W. I. Newman, W. M. Kaula and J. M. Hyman, "Dynamical Evolution of Planetesimals in Outer Solar System," *Icarus*, Vol. 140, No. 2, 1999, pp. 341-352. doi:10.1006/icar.1999.6146

[3] K. Heun, "Neue Methode zur Approximativen Integration der Differentialgleichungen einer Unabhängigen Veränderlichen," *Mathematical Physics*, Vol. 45, 1900, pp. 23-38.

[4] M. W. Kutta, "Beitrag zur Näherungsweisen Integration totaler Differentialgleichungen," *Mathematical Physics*, Vol. 46, 1901, pp. 435-453.

[5] F. T. Krogh, "A Variable Step Variable Order Multistep Methods for Ordinary Differential Equations," *Information Processing Letters*, Vol. 68, 1969, pp. 194-199.

[6] E. J. Nyström, "Über die Numerische Integration von Differentialgleichungen," *Acta Societas Scientiarum Fennicae*, Vol. 50, No. 13, 1925, pp. 1-54.

[7] C. Störmer, "Sur les Trajectoires des Corpuscles Électrisés," *Acta Societas Scientiarum Fennicae*, Vol. 24, 1907, pp. 221-247.

[8] D. Brouwer, "On the Accumulation of Errors in Numerical Integration," *Astronomical Journal*, Vol. 46, No. 1072, 1937, pp. 149-153. doi:10.1086/105423

[9] K. R. Grazier, W. I. Newman, J. M. Hyman and P. W. Sharp, "Long Simulations of the Outer Solar System: Brouwer's Law and Chaos," In: R. May and A. J. Roberts, Eds., *Proceedings of* 12*th Computational Techniques and Applications Conference* CTAC-2004, *ANZIAM Journal*, Vol. 46, 2005, pp. C1086-C1103.

[10] E. Hairer, R. I. McLachlan and A. Razakarivony, "Achieving Brouwer's Law with Implicit Runge-Kutta Methods," *BIT Numerical Mathematics*, Vol. 48, No. 2, 2008, pp. 231-243. doi:10.1007/s10543-008-0170-3

[11] W. H. Enright, D. J. Higham, B. Owren and P. W. Sharp, "A Survey of the Explicit Runge-Kutta Method," Technical Report, 291/94, Department of Computer Science, University of Toronto, Toronto, 1994.

[12] J. Dormand, M. E. A. El-Mikkawy and P. Prince, "Higher

Order Embedded Runge-Kutta-Nyström Formulae," *IMA Journal of Numerical Analysis*, Vol. 7, No. 4, 1987, pp. 423-430. doi:10.1093/imanum/7.4.423

[13] L. F. Shampine and M. K. Gordon, "Computer Solution of Ordinary Differential Equations," W. H. Freeman, San Francisco, 1975.

[14] E. Hairer, S. P. Nørsett and G. Wanner, "Solving Ordinary Differential Equations I: Nonstiff Problems," Springer-Verlag, Berlin, 1987.

[15] K. R. Grazier, "The Stability of Planetesimal Niches in the Outer Solar System: A Numerical Investigation," PhD Thesis, University of California, 1997.

