CONTINUOUS APPROXIMATIONS FOR LONG-TERM NUMERICAL SIMULATIONS OF THE SOLAR SYSTEM

S. Rehman
Department of Mathematics,
University of Engineering and Technology, Lahore, Pakistan,
srehman@uet.edu.pk

(Received 16 January 2021; in final form 28 July 2021; accepted 26 August 2021; published online 30 September 2021)

We present and analyse the performance of different combinations of four higher-order numerical integrators and up to nine interpolation schemes applied to the problem involving the Sun and four Gas-giants (outer planets), namely, Jupiter, Saturn, Uranus, and Neptune. The Hermite interpolation schemes obtained by one, two, and three time-step and interpolants for ODEX2 and ERKN integrators are considered in this paper. The interpolants are a special example of an interpolation scheme, which produce an approximation that is continuous across one step and across the complete interval of integration. The interpolants are quite expensive in comparison with the other interpolation schemes. Therefore, one of the objectives of this paper is to investigate the possibilities of replacing the interpolants of certain integrators by other interpolation schemes, perhaps at a cost of a little bit of accuracy. The experiments are performed to examine the error growth in the positions, velocities, and relative error in energy and angular momentum using different combinations of integrators and interpolation schemes over a long interval of integration, as long as 100 million years for the Jovian problem with local error tolerances ranging from 10^{-16} to 10^{-08}.

Key words: Jovian problem, interpolation schemes, long-term simulations.

DOI: https://doi.org/10.30970/jps.25.3901

I. OVERVIEW

The interpolation schemes play a key role in N-body simulations, in-particular, when detecting close-encounters between a massive body and a test particle in the Solar System dynamics. It is possible that a close-encounter occurs within the time-step and not at the end points. Such a close-encounter would not be detected if we only have numerical approximations at the end points. One way of detecting such close-encounters is to generate numerical approximations within the integration time-step by using a far smaller time-step when the test particle is near the massive body compared with when it is far from the massive body. In practice, this approach would be inefficient because all bodies, including those that are not undergoing close-encounters, would be integrated with the smaller time-step, resulting in a possibly tremendous increase in CPU-time. Thus, an alternative approach is needed to avoid this inefficiency. The numerical approximation at the mesh points may be extended to a continuous approximation, which provides the numerical solution at any point t, for t within the time-step interval. Large numbers of numerical integrators and associated interpolation schemes for performing N-body simulations have been developed and implemented; see, for example, [9, 18–22].

In the next section, we illustrate the physical formulation of the Jovian problem, integrators and interpolation schemes applied to the Jovian problem, different types of errors, estimation of the maximum global error by sampling the norm with different sets of data points. In Section III, we present continuous approximations of variable-step-size integrators, and a schematic that presents the classification of the errors as a local segment of the numerical solution. Numerical testing is done in Section IV involving comparisons of combinations of different integrators and interpolation schemes, and overall summary is presented in Section V. All computer programs are written in FORTRAN and we have used Matlab to analyse the results. The graphs are smoothed with the Matlab filter command.

II. BACKGROUND

The Jovian problem models the orbital motion of the Sun and four Gas giants (outer planets), namely, Jupiter, Saturn, Uranus, and Neptune, interacting with one another through Newtonian forces [23]. The shortest orbital period for the Jovian problem is 4331 days (Jupiter). The Jovian planets collectively drive much of the dynamics of our Solar System. Therefore, the Jovian problem is frequently used in numerical approximations, inparticular, when simulations are performed over a long interval of integration. These longterm numerical simulations provided more insight into the Solar System dynamics, which went further than that given by analytic theories. Let $r_i = [x_i, y_i, z_i]^T, i = 1, \ldots, 5$, denote the position of the ith body of the Jovian problem in a three-dimensional Cartesian coordinate system with the origin at the barycentre of the bodies. The equations of motion for the ith body of the Jovian problem can be written as

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

where $||.||_2$ is the L_2-norm, and μ_j denotes the gravitational constant G times the mass m_j of the jth body, i.e., $\mu_j = Gm_j$, distance is expressed in astronomical units, time in Earth days, whereas the mass m_j in Solar mass.

For each body of the Jovian problem we have a 2nd-order differential equation for the x-, y-, and z-components, giving us 15 2nd-order differential equations in total. For efficiency purposes, we use the symmetry of interactions for evaluating the acceleration for the Jovian problem. If we consider the individual terms in the summation, then it can be observed that

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

Once this term for r_j is obtained, we can easily update the acceleration for the 2nd body by using symmetry. It has been observed that using symmetry, the subroutine for evaluating the force term reduces to approximately half of the CPU-time.

Explicit Runge–Kutta–Nyström (ERKN) methods for the numerical approximation of 2nd-order differential equations were proposed by E. J. Nyström in 1925 [12]. ERKN methods reduce the computational cost considerably, compared to explicit Runge–Kutta (ERK) methods applied to the corresponding equivalent system of first-order differential equations. For example, an order-five ERKN method requires only four function evaluations per integration time-step, whereas an ERK method of the same order requires at least six function evaluations [10]. The efficiency of an ERKN method depends on the technique for controlling the error in the numerical solutions. An adaptive step-size technique is one of the possible options of controlling the error that permit control of the estimated local error. A pair of different orders formulae is implemented in such a manner that the function evaluations of both methods are identical. Usually, the numerical approximation is performed by a higher-order method and the error is obtained by the lower-order method to gain maximum efficiency. Here, we use two ERKN integrators: a 9-stage *ERKN689*, 6–8 FSAL pair and a 17-stage *ERKN101217*, 10–12 non-FSAL pair [2].

Extrapolation provides a strong means of accelerating the convergence of solutions that arise from discretization methods, and also have strong connections with, for example, projection methods, continued functions and Padé approximations. For the direct numerical approximations of 2nd-order systems of differential equations, Hairer et al. [10] developed an extrapolation code *ODEX2*. This code is based on the explicit midpoint rule along with order selection and a step size control technique. The extrapolation code *ODEX2* is efficient for all tolerances, in-particular for high precision.

Störmer's methods are an important class of numerical methods for the numerical approximation of systems of 2nd-order differential equations [14]. Störmer's methods have long been implemented for long-term numerical simulations of the Solar System dynamics [9]. Grazier [9] suggested the 13th-order, fixed step-size Störmer method that uses backward differences in summed form, summing from the highest to the lowest differences. For the Jovian problem, the numerical testing in [8] depicts that the error in energy and the phase error grow as $t^{1/2}$ and $t^{3/2}$, respectively, to within numerical uncertainty when the step size is approximately 4 days of Jupiter's orbital period. This particular selection of the step-size guarantees that the local truncation error is well below machine precision. In this work, we use the order-13, fixed step-size Störmer method and associate to it as *$\bar{S}$-13* integrator.

For the continuous approximation of the Jovian problem, the interpolation schemes used in this paper are: one-step (cubic and quintic Hermite interpolation schemes), two-step, and three-step Hermite interpolation schemes. The cubic Hermite interpolation polynomial is of degree 3, whereas the quintic, two-step, and three-step are of degrees 5, 8, and 11, respectively. For one time-step, the cubic Hermite interpolation polynomial interpolates the data (t_{n-i}, y_{n-i}) and (t_{n-i}, y'_{n-i}) at time t_{n-i}, for $i = 1$ and 0 and can be expressed as

$$\begin{aligned} P_3(t) &= (\tau-1)^2(2\tau+1)y_{n-1} + (\tau-1)^2\tau H y'_{n-1} \\ &+ \tau^2(3-2\tau)y_n + \tau^2(\tau-1)Hy'_n \ , \end{aligned}$$

where, $H = t_n - t_{n-1}$ and $\tau = (t - t_{n-1})/H$. Since the values of y and y' are interpolated at both ends of each time-step, the piecewise defined approximation obtained from the cubic polynomial is continuous and has a continuous first derivative.

The quintic polynomial also interpolates the data involving position, velocity, and acceleration at time t_{n-i}, $i = 1, 0$. The values of y, y', and y'' are interpolated at both ends of each time-step. Therefore, the piecewise defined approximation obtained from the quintic polynomial is continuous and has continuous first- and second-derivatives. Similarly, the two-step polynomial interpolates the data involving position, velocity, and acceleration at time t_{n-i}, $i = 2, 1, 0$; whereas, the three-step interpolation polynomial interpolates the data involving position, velocity, and acceleration at time t_{n-i}, $i = 3, 2, 1, 0$. For comprehensive details of the one-step, two-step, and three-step interpolation schemes, we refer to [17]. Solution interpolants for *ODEX2* integrator and ERKN integrators are discussed in Sections III A and III B, respectively.

Now, we define different types of errors used in this paper. We use the notation $y_{\text{true}}(t)$ to denote the true solution and $y_{\text{num}}(t)$ to denote the approximate solution. The difference is monitored time-wise by considering the global error in y as

$$y_{\text{num}}(t) - y_{\text{true}}(t). \tag{2}$$

The norm of the global error in y at time t is then

$$||y_{\text{num}}(t) - y_{\text{true}}(t)||_2. \tag{3}$$

The calculation of the global error is discussed later. A large number of integrators, for example, Runge–Kutta–Nyström [12], and Störmer [14] can be used to find $y_{\text{num}}(t)$ for $t \geq 0$. This leads to numerical solutions $y_n = y_{\text{num}}(t_n)$ and $y'_n = y'_{\text{num}}(t_n)$ at times $t_n = t_0 + nh$, $n = 1, 2, \ldots$, where h can depend on n. The approximation over the continuous time interval can then be computed using (local) interpolation.

From one time-step to the next, the local problem is solved, which is defined as

$$u_n'' = f(t, u_n), \quad u_n(t_{n-1}) = y_{n-1},$$
$$u_n'(t_{n-1}) = y_{n-1}', \quad t \in [t_{n-1}, t_n], \tag{4}$$

where $u_n(t)$ is the true local solution on the nth interval. The initial conditions y_{n-1} and y_{n-1}' are the values of the approximate solution at the end of the $(n-1)^{\text{st}}$ step. Hence, for the local problem, the DEs are the same as for the original problem, while the initial conditions depend upon the numerically approximated solution. Because, in general, the approximation y_{n-1} is not equal to the true solution $y_{\text{true}}(t_{n-1})$, there is a difference between the error obtained in this local step and the global error. The norm of the local error is defined as

$$||y_{\text{num}}(t) - u_n(t)||_2, \tag{5}$$

where $t \in [t_{n-1}, t_n]$. In general, the global error cannot be calculated because the true solution is not known.
The estimated global error in the position over each time-step, n, is defined in terms of an accurate numerical solution (computed in quadruple precision), $y_{\text{ref}}(t)$, and an underlying polynomial approximation is introduced at each step, $P_n(t)$. We then estimate global error for $t \in [t_{n-1}, t_n]$ by evaluating $P_n(t) - y_{\text{ref}}(t)$ at k evenly spaced sample points over the step. A more detailed description of $y_{\text{ref}}(t)$, k, $P_n(t)$, and the definition of the estimated global error in the position is presented later in this section.

The global error in the position for all $t \in [t_0, t_f]$, where t_f is the pre-specified final value of t in the integration, is more difficult to define and estimate, because we need a continuous approximation to $y(t)$, and not just the discrete values $y_1, y_2, \ldots$. We obtain the continuous approximation by first forming a polynomial $P_n(t)$ that approximates the local solution on the interval $[t_{n-1}, t_n]$. The numerical approximation is then the piecewise-defined function

$$y_{\text{num}}(t) = \begin{cases} P_1(t), & t_0 \le t \le t_1, \\ P_2(t), & t_1 \le t \le t_2, \\ \vdots & \\ P_n(t), & t_{n-1} \le t \le t_n, \\ \vdots & \end{cases}$$

where we have assumed $P_{n-1}(t_{n-1}) = P_n(t_{n-1})$, an assumption that holds throughout the paper. The corresponding definition of $y_{\text{num}}'(t)$ is

$$y_{\text{num}}'(t) = \begin{cases} \hat{P}_1(t), & t_0 \le t \le t_1, \\ \hat{P}_2(t), & t_1 \le t \le t_2, \\ \vdots & \\ \hat{P}_n(t), & t_{n-1} \le t \le t_n, \\ \vdots & \end{cases}$$

where $\hat{P}_n(t)$ not necessarily is the derivative of $P_n(t)$. The polynomials $P_n(t)$ and $\hat{P}_n(t)$ are commonly called local interpolants and can be of many types. The main requirements of the local interpolants are that they be sufficiently accurate and have sufficient continuity.
We write the norm of the estimated global error in the position at time t as

$$E_r(t) = ||y_{\text{num}}(t) - y_{\text{ref}}(t)||_2, \tag{6}$$

and the norm of the estimated end-point global error as

$$E_{r,\text{end}}(t_f) = ||y_{\text{num}}(t_f) - y_{\text{ref}}(t_f)||_2, \tag{7}$$

where t_f is the time at the end-point. The maximum of the norm of the estimated global error in position, loosely referred to as the maximum global error, is defined as the maximum norm of $E_{r,\max}(t)$ evaluated at k evenly spaced sample points over each integration time-step. That is, we evaluate the maximum global error by introducing an accurate approximation $y_{\text{ref}}(t)$ to $y(t)$, sampling $||y_{\text{num}}(t) - y_{\text{ref}}(t)||_2$ over each time-step, and using the maximum over all sampled values to be the estimate of the maximum of $||y_{\text{num}}(t) - y(t)||_2$.
In way a similar to that for the position, the norm of the estimated global error in velocity at time t and the maximum of this norm over the interval $[t_0, t_f]$ are defined as

$$E_v(t) = ||y_{\text{num}}'(t) - y_{\text{ref}}'(t)||_2,$$
$$E_{v,\max}(t) = \max_{t \in [t_0, t_f]} ||y_{\text{num}}'(t) - y_{\text{ref}}'(t)||_2,$$

where $y_{\text{num}}'(t)$ and $y_{\text{ref}}'(t)$ are vectors at any time t of the derivatives to the numerical and reference solutions, respectively.

We store the pieces of information, like positions and times, in separate files. Meanwhile, in a post-processing program, we estimate the maximum global error by sampling the norm at k evenly spaced data points on every sub-interval $[t_{n-1}, t_n]$ with respect to the reference solution, which we obtain at these stored values of time by forcing the integrator to hit the sample points, and then taking the maximum of these norms. To see the effect of different values of k on the estimated maximum global error in position, we performed experiments using the combination of *ERKN101217* and the 29-stage interpolant, which are described in Section III A. The integration was performed over one million years of integration for the Jovian problem with local error tolerances 10^{-08}, 10^{-10}, 10^{-12}, 10^{-14}, and 10^{-16}. We found that $k = 10$ led to an estimated maximum global error that was sufficiently accurate for our work; this is illustrated by the results in Table 1. The rows labelled $k = 50$ and $k = 100$ list the percentage changes in the estimated maximum global error when compared with the values in row labelled $k = 10$.

We observe from Table 1 that the data for $TOL = 10^{-08}$ shows a percentage change of zero. In this case, the estimated maximum global error occurs at t_f and this is possible with N-body simulations; if this occurs, the percentage difference between different values in k

would be zero. We repeated the same set of experiments with other combinations. For example, when using the integrator *ODEX2* with its interpolant, described in Section III B, an increase of k from 10 to 100 changed the estimated maximum global error by not more than 1%. Throughout the remainder of this paper, all maximum global errors are estimated by sampling at 10 evenly spaced data points at every time-step.

TOL	10^{-08}	10^{-10}	10^{-12}	10^{-14}	10^{-16}
$k=10$	4.70×10^{-1}	1.63×10^{-3}	4.82×10^{-5}	2.42×10^{-5}	8.80×10^{-6}
$k=50$	0 %	3.4×10^{-2} %	3.9×10^{-3} %	1.0×10^{-2} %	5.2×10^{-5} %
$k=100$	0 %	3.5×10^{-2} %	3.9×10^{-3} %	1.0×10^{-2} %	4.2×10^{-5} %

Table 1. The values of estimated maximum global error using *ERKN101217* with 29-stage interpolant obtained with different values of k over one million years of integration for the Jovian problem with local error tolerances 10^{-08}, 10^{-10}, 10^{-12}, 10^{-14}, and 10^{-16}. The row labelled $k=10$ shows errors, whereas the rows labelled $k=50$ and $k=100$ list the percentage changes.

Usually, physical systems have conserved quantities, such as, total energy, total angular momentum, position and velocity of the center of mass of the bodies. Normally, these physical quantities are not conserved exactly by the numerical approximation and this deviation provides an insight about the accuracy of the numerical approximation. The total energy for a system of N bodies interacting with one another through Newtonian forces is defined as

$$H(t) = \frac{1}{2}\sum_{i=1}^{N} m_i(v_{i,\mathrm{num}}(t)\cdot v_{i,\mathrm{num}}(t)) - \sum_{j=1}^{N-1}\sum_{i=j+1}^{N} G\frac{m_i m_j}{d_{ij}(t)},$$

where G is the gravitational constant, m_i the mass of the ith body, $v_{i,\mathrm{num}}(t)$ the numerical approximation to velocity of the ith body (components $3i-2$ to $3i$ of $y'_{\mathrm{num}}(t)$), and $d_{ij}(t) = ||r_{i,\mathrm{num}}(t) - r_{j,\mathrm{num}}(t)||_2$ is the distance between the ith and j^{th} bodies. Here, $r_{i,\mathrm{num}}(t)$ is formed by the components $3i-2$ to $3i$ of $y_{\mathrm{num}}(t)$ and $r_{j,\mathrm{num}}(t)$ by the components $3j-2$ to $3j$ of $y_{\mathrm{num}}(t)$, which represent the numerical approximations to the positions of the ith and j^{th} bodies, respectively. The relative error in energy is defined as

$$H_{\mathrm{rel}}(t) = \left|\frac{H_0 - H(t)}{H_0}\right|,$$

where H_0 is the total energy at $t = t_0$. The value of $\mu = Gm$ is usually known more accurately than the value of m. Therefore, we use

$$H_{\mathrm{rel}}(t) = \left|\frac{GH_0 - GH(t)}{GH_0}\right|.$$

The total angular momentum is defined as

$$L(t) = \sum_{i=1}^{N} m_i r_{i,\mathrm{num}}(t) \times v_{i,\mathrm{num}}(t).$$

We define the relative error of the angular momentum as

$$L_{\mathrm{rel}}(t) = \frac{||L_0 - L(t)||_2}{||L_0||_2},$$

where L_0 is the angular momentum at the initial time $t = t_0$.

III. CONTINUOUS APPROXIMATIONS

The well-known one-step methods for the numerical approximation of a system of ordinary differential equations are normally formulated such that numerical approximations are produced on mesh points $t_0 < t_1 < t_2 < \ldots < t_f$, which are determined by the step-size selection strategy of these numerical methods. As soon as a numerical approximation of $y_{\mathrm{true}}(t)$ is required where t is not at one of the mesh points, the numerical approximation must be extended to a continuous approximation. Pioneering work on continuous approximations for Runge–Kutta methods has appeared in the literature [11], and also for explicit methods [13]. This work was extended by others; see, for example, [1, 3–6, 10, 15, 16].

Figure 1 shows a local segment of the numerical solution y_{num} over the time interval $[t_{n-1}, t_n]$. This solution is obtained by interpolation over the data calculated from the local problem (4). The local solution to (4) has a local error that is 0 at $t = t_{n-1}$. In contrast, the global error is not 0 for any $t \in [t_{n-1}, t_n]$. There are two contributions to the global error. First, the interpolation polynomial is not formed using true data. Rather, it is using an approximation at the end of the time-step. Second, even if it is true data then there is an interpolation error at the intermediate times $t \in [t_{n-1}, t_n]$. If we denote the estimated global error in an interpolated value on $[t_{n-1}, t_n]$ by $E_{\mathrm{r,int}}(t)$, then the relative error with respect to the estimated end-point global error is

$$\begin{aligned} R_{\mathrm{err}}(t) &= \frac{\text{Estimated global error in an interpolated value on } [t_{n-1}, t_n]}{\text{Estimated global error in the numerical solution at } t_n}, \\ &= \max_{t\in[t_{n-1},t_n]} \frac{E_{\mathrm{r,int}}(t)}{E_{\mathrm{r,end}}(t_n)}. \end{aligned} \tag{8}$$

If the integration error dominates the interpolation error, then the relative error $R_{\mathrm{err}}(t)$ tends to 1 as $t \mapsto \infty$.

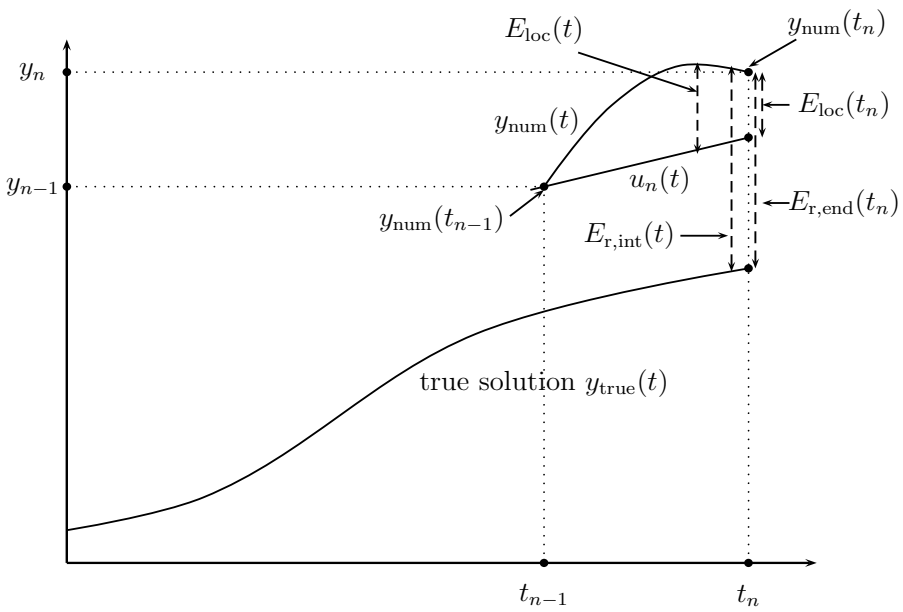

Fig. 1. Classification of the errors when using a numerical integrator together with an interpolation polynomial

A. Continuous approximation with embedded RKN methods

Dormand and Prince [3] and then Baker *et al.* [1] constructed a continuous approximation for embedded RKN pairs, where a third RKN process of order p^* is used for the numerical approximations of $y(t_{n-1+\alpha})$ and $y'(t_{n-1+\alpha})$ with α typically in (0, 1]. These approximations are usually called interpolants and are denoted by $y^*_{n-1+\alpha}$ and $y'^*_{n-1+\alpha}$:

$$\begin{aligned} y^*_{n-1+\alpha} &= y_{n-1} + \alpha h_{n-1} y'_{n-1} + \alpha^2 h^2_{n-1} \sum_{i=1}^{s^*} b^*_i(\alpha) k^*_i, \\ y'^*_{n-1+\alpha} &= y'_{n-1} + \alpha h_{n-1} \sum_{i=1}^{s^*} b'^*_i(\alpha) k^*_i \end{aligned} \tag{9}$$

where,

$$\begin{aligned} k^*_i &= f(t_{n-1} + c^*_i \alpha h_{n-1},\ y_{n-1} + c^*_i \alpha h_{n-1} y'_{n-1} \\ &+ \alpha^2 h^2_{n-1} \sum_{j=1}^{i-1} a^*_{ij} k^*_j), \quad i = 1, \ldots, s^*. \end{aligned}$$

If $s^* = s$, then no extra function evaluations are required for $y^*_{n-1+\alpha}$ and $y'^*_{n-1+\alpha}$. As before, the prime in b'^*_i is purely notational and does not involve the differentiation operator. The approximation to $y_{\rm num}(t)$ is continuous and has continuous first- and second-derivatives, because it interpolates y, y', and y'' at both ends of each time-step. For ERKN methods, the derivative interpolants can be derived separately, and not as the derivative to the solution interpolants themselves.

For *ERKN101217* integrator, we have used 23-stage, 26-stage, and 29-stage interpolants with $p^* = 10$, $p^* = 11$, and $p^* = 12$, respectively. The coefficients for these interpolants are freely available on-line [1]. For *ERKN689* integrator, we have used a 12-stage interpolant with $p^* = 8$. Note that the coefficients for the interpolant for y and y' of *ERKN689* integrator are provided (private communication) by P. W. Sharp.

B. Continuous approximation with *ODEX2*

The *ODEX2* integrator has a solution interpolant. We added a derivative interpolant by differentiating the solution interpolant. The solution interpolant provides an approximation to the $i^{\rm th}$-component of the solution at time t. The polynomial $P_\mu(\alpha)$ for the solution interpolant can be written as

$$\begin{aligned} P_\mu(\alpha) &= y_{(i)} + \alpha y_{(M+i)} + \alpha(1-\alpha) y_{(2M+i)} + \alpha^2(1-\alpha) y_{(3M+i)} + \alpha^3(1-\alpha)^2 y_{(4M+i)} \\ &+ \alpha^2(1-\alpha)^3 y_{(5M+i)} + (\alpha - \alpha^2)^3 \sum_{j=0}^{k} \frac{(\alpha_1)^j}{j!} y_{(M(5+j)+i)}, \end{aligned}$$

where $\alpha_1 = (t - t_{n-1})/h - 0.5$ and M is the number of ordinary differential equations. The degree of $P_\mu(\alpha)$ is $\mu + 4$, with $-1 \le \mu \le 2k$, where k is related to the integration order.

The polynomial used for the continuous approximation of the derivative provides an approximation to the ith-component of the derivative at time t. The order of this derivative interpolant is one less than $P_\mu(\alpha)$ and can be written as

$$P'_\mu(\alpha) = \frac{1}{h}\Big[y_{(M+i)} + (1-2\alpha)y_{(2M+i)} + (2\alpha-3\alpha^2)y_{(3M+i)} + \alpha^2(1-\alpha)(3-5\alpha)y_{(4M+i)} + \alpha(1-\alpha)^2(2-5\alpha)y_{(5M+i)} + (\alpha-\alpha^2)^3\sum_{j=1}^{k}\frac{(\alpha_1)^{j-1}}{(j-1)!}y_{(M(5+j)+i)} + 3(\alpha-\alpha^2)^2(1-2\alpha)\sum_{j=0}^{k}\frac{(\alpha_1)^j}{j!}y_{(M(5+j)+i)}\Big].$$

IV. NUMERICAL TESTING FOR LONG-TERM SIMULATIONS

First of all, we discuss error growth in position and velocity by using the appropriate combinations integrators: *ODEX2*, *ERKN689*, and *ERKN101217*, and interpolation schemes over one million years for the Jovian problem.

We have obtained the reference solution in quadruple-precision using the combination of *ERKN101217* integrator and the 29-stage interpolant with $TOL = 10^{-18}$. To quantify this particular choice for the reference solution, we also integrated the Jovian problem using the combination of *ERKN101217* and the 29-stage interpolant with $TOL = 10^{-20}$. It has been observed that the maximum difference between positions and velocities of these two solutions is no more than 4.6×10^{-13}. It has also been observed that in most cases, the maximum of the global error occurs at the end-point of the integration. We also integrated the Jovian problem in quadruple-precision using the combination of *ERKN689* and the 12-stage interpolant with the tolerance $TOL = 10^{-18}$ and observed that the maximum difference with the solution for the combination of *ERKN101217* and the 29-stage interpolant with $TOL = 10^{-18}$ is no more than 5.1×10^{-13}. This suggests that the combination of *ERKN101217* and the 29-stage interpolant with $TOL = 10^{-18}$ is fairly accurate to acquire the reference solution.

We have performed numerical experiments to observe the unweighted L_2-norm of the estimated maximum global error in the position as a function of tolerance with three different combination; *ODEX2* and its built-in interpolant, *ERKN689* and the 12-stage interpolant, and *ERKN101217* and the 29-stage interpolant over one million years for the Jovian problem. It has been observed that the maximum global error obtained with the combination of *ODEX2* and its built-in interpolant is ranging from 8.0×10^{-5} to 1.1×10^{2}. The best observed accuracy 8.0×10^{-5} is obtained with $TOL = 10^{-16}$ and the minimum accuracy 1.1×10^{2} with $TOL = 10^{-08}$. The combination of *ERKN101217* and the 29-stage interpolant is an accuracy that ranges from 8.8×10^{-6} to 4.7×10^{-1}. The best observed accuracy is obtained with $TOL = 10^{-16}$ and the minimum accuracy with $TOL = 10^{-08}$. Similarly, the combination of *ERKN689* and the 12-stage interpolant has an accuracy ranging from 9.8×10^{-7} to 4.4×10^{-2}, with the best observed accuracy at $TOL = 10^{-14}$ and the minimum accuracy is obtained at $TOL = 10^{-08}$.

To test the analysis that led to the definition of the error R_{err} (8), we integrate the Jovian problem over one million years using all the integrators and interpolation schemes described so far. On every accepted integration time-step, we obtain the L_2-norm of H_{rel} and L_{rel} at 10 evenly spaced values of time. The maximum of these 10 error values is taken as the maximum error (care has been taken to make efficient use of storage) on the integration time-step. The $\bar{S}$*-13* integrator uses a fixed step-size of 4 days, and we set the tolerances to 10^{-16}, 10^{-11}, and 10^{-10} for *ODEX2*, *ERKN101217*, and *ERKN689*, respectively (the variation of tolerance is subject to achieving a maximum global error of 10^{-4} with variable-step-size integrators).

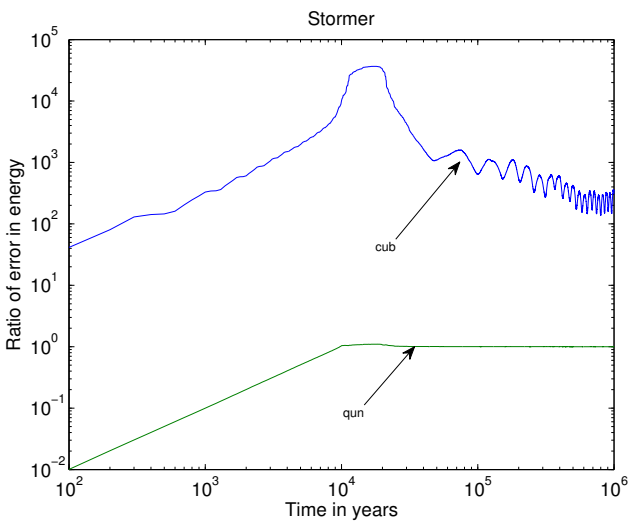

Fig. 2. The ratio of the maximum H_{rel} at the interior points to the end-point (of a step) H_{rel} for the $\bar{S}$*-13* integrator using cubic and quintic Hermite interpolations over one million years for the Jovian problem

Figure 2 depicts the ratio of the maximum H_{rel} at the interior points to the end-point H_{rel} for the $\bar{S}$*-13* integrator using cubic and quintic interpolations. We observe from the plot for the cubic Hermite interpolation polynomial that R_{err} appears bounded from above by 5×10^4 for small t and then gradually decreases to no more than 6×10^2 at 10^6 years. The large value

of R_{err} may be due to the fact that the interpolation polynomial used for the velocity components is only of order 2. The plot for the quintic Hermite interpolation polynomial shows that the relative error in energy is approximately 1 for the entire interval of integration. This verifies the result of Grazier et al. [7] that, when the time-step is chosen so that the $\bar{S}$-*13* methods satisfy Brouwer's law, the one-step quintic Hermite interpolation is sufficiently accurate. A similar behaviour with cubic and quintic polynomials is observed for the relative error in angular momentum.

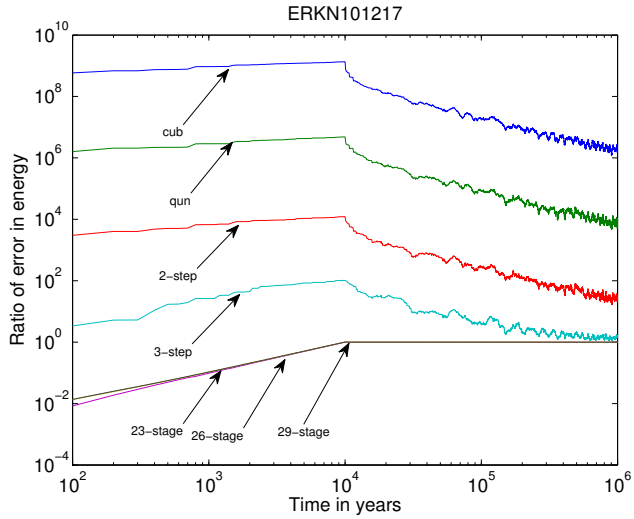

Fig. 3. The ratio of the maximum H_{rel} at the interior points to the end-point H_{rel} for *ERKN101217* using cubic, quintic, 2-step, 3-step interpolations, and 23-stage, 26-stage, 29-stage interpolants over one million years of integration for the Jovian problem with local error tolerance of 10^{-11}

Figure 3 represents the ratio of the maximum H_{rel} at the interior points to the end-point H_{rel} for *ERKN101217* using cubic, quintic, 2-step and 3-step interpolations [17], which are of degrees 3, 5, 8, and 11, respectively; and 3 interpolants of 23, 26, and 29 stages, having orders 10, 11, and 12, respectively. We observe for the low-order interpolation schemes, i.e., cubic, quintic, and two-step, that R_{err} in energy does not decrease to 1 at the end of 10^6 years. With all three interpolants the R_{err} is approximately 1 for the entire interval of integration. We also observe that the 3-step interpolation polynomial achieves the same accuracy (with R_{err} in energy approximately 1) at the end of the integration, although initially it is dominated by the interpolation error. Therefore, until about 10^4 years, it is better to use one of the interpolants if the smallest error is required. However, if the required integration time is large enough, the integration error starts dominating and then it is better to switch to a 3-step interpolation polynomial because this is going to save on CPU-time and does not sacrifice accuracy.

We also performed experiments measuring R_{err} for the energy and angular momentum for *ERKN689* using the cubic, quintic, and two-step Hermite interpolation schemes, and together with its 12-stage interpolant; and for *ODEX2* using cubic, quintic, two-step and three-step interpolations, and together with the interpolant. For *ERKN689*, the two-step interpolation scheme and the 12-stage interpolant achieved a ratio of approximately 1. For *ODEX2*, the only continuous approximation that achieved the same accuracy was the interpolant and its derivative that comes with *ODEX2*.

To observe the behaviour of conserved quantities like energy and angular momentum using different interpolation schemes with different integrators, we performed many experiments. All these experiments were done for the Jovian problem over up to 10^8 years. One of the strong observations about all of these experiments that applies except when using the quintic Hermite interpolation scheme with the $\bar{S}$-*13* integrator is as follows: if a low-order interpolation scheme is used with a high-order integrator then the interpolation error at a particular time will dominate the integration error and the total error will not increase with t. This is because the interpolation error does not grow with time. For example, if a 12-th order integrator is used with cubic Hermite interpolation, then the total error is given by

$$\begin{aligned}\text{Total error} &= Ct^{3/2}h^{12} + Dh^4 \\ &\approx Dh^4, \qquad C, D = \text{Constant},\end{aligned}$$

provided t is not too large and C is not significantly larger than D. On the other hand, if the interpolation polynomial and the integrator are of the same order, then the total error will increase with the passage of time.

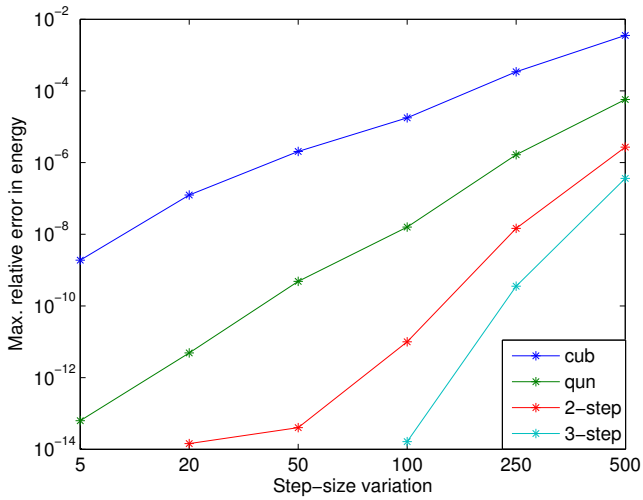

Fig. 4. The maximum H_{rel} for the cubic, quintic, 2-step, and 3-step interpolations with different step-size sequences over one period of Jupiter for the Jovian problem

We now want to observe the effect of the step-size on H_{rel} and L_{rel} while performing long-term simulations using different interpolation schemes with different integrators. Therefore, we set up an experiment to anticipate the behaviour of this relative error; see Figure 4. This experiment provides an insight about the interpolation error in energy and angular momentum. We integrate the Jovian problem using the *ERKN101217* integrator over a time interval that equals one period of

Jupiter. We obtain a reference solution by integrating in quadruple precision with $TOL = 10^{-18}$ to eliminate the possible effect of integration error. We run six experiments, where we record position, velocity and acceleration at the end of every 5, 20, 50, 100, 250, and 500 days, by forcing the integrator to hit these time-points. The period of Jupiter is not divisible by 5, 10, 50, 100, 250, and 500 and hence the last step of each integration was shorter than the previous steps. The selection of such step-size sequences is done in accordance with the average step-sizes taken by different integrators when integrating the Jovian problem. For example, over 10^6 years, *$\bar{S}$-13* takes a time-step of 4 days, while on average, over the same time interval, *ERKN689* with $TOL = 10^{-10}$, *ERKN101217* with $TOL = 10^{-11}$ and *ODEX2* with $TOL = 10^{-16}$ take approximately 65, 290 and 260 days, respectively.

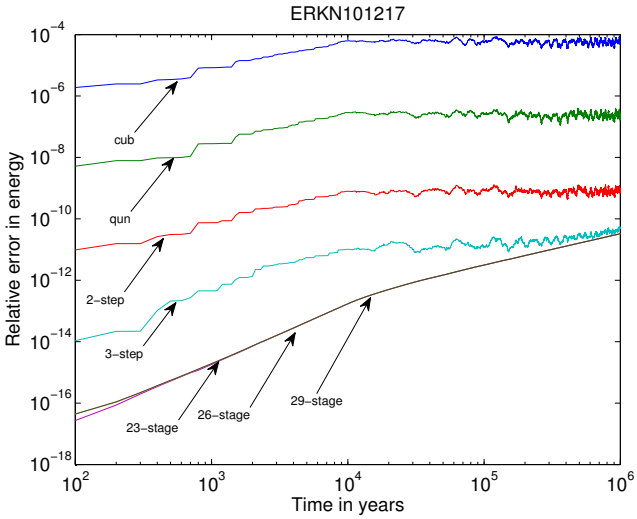

Fig. 5. The $H_{\rm rel}$ for *ERKN101217* using cubic, quintic, 2-step, 3-step interpolations, and 23-stage, 26-stage, 29-stage interpolants over one million years of integration for the Jovian problem with the local error tolerance of 10^{-11}

For different step-size sequences, we use appropriate interpolation schemes; for example, as discussed earlier, the 3-step interpolation scheme needs not to be used for the *$\bar{S}$-13* integrator with a step-size of 4 days, since the quintic Hermite interpolation is sufficient. Then on each step-size, we use interpolation and sample the solution at 10 equally-spaced data points, from which we determine the local maxima for $H_{\rm rel}$ and $L_{\rm rel}$. For each interpolation scheme, we observe that there is a considerable variation in $H_{\rm rel}$ and $L_{\rm rel}$ as a function of the step-size. For example, the biggest variation of approximately 9 orders of magnitude for $H_{\rm rel}$ was observed with the quintic Hermite interpolation scheme over the range of 5 to 500 days step-size sequences. This particular example has a reasonably good agreement with the expected difference of approximately 10 orders of magnitude, because the expression for the relative error in energy is dominated by the velocity term of order 5 with a quintic interpolation. This means that, by using expression $\alpha h^{p+1} \frac{y_{\rm true}^{(p+1)}(\xi)}{(p+1)!}$ for the interpolation error, the expected difference of 5 and 500 days step-size sequences is $(500/5)^5 = 10^{10}$.

Figure 5 contains the graphs of the relative error in energy for *ERKN101217* when used with the cubic, quintic (one-step), 2-step and 3-step interpolation polynomials, of degrees 3, 5, 8 and 11, respectively; and the three interpolants of 23, 26, and 29 stages, having orders 10, 11, and 12, respectively. The *ERKN101217* integrator for an integration of 10^6 years using $TOL = 10^{-11}$ requires, on average, a step-size of approximately 290 days. We find that the total error for the cubic, quintic, and 2-step interpolations does not increase with t. It remains pretty much constant, as was expected, because the interpolation error dominates the integration error.

The interpolation errors at the end of the integration have a reasonably good agreement with the results obtained in Figure 4. With all three interpolants, i.e., 23-stage, 26-stage and 29-stage, we have obtained an accuracy of approximately 10^{-11}. With the 3-step interpolation scheme, which is cheaper than all three interpolants, the accuracy is about the same at the end of the integration, although initially it is dominated by the interpolation error. Therefore, it is better to use an interpolant if the best accuracy is required. However, if the time interval is large enough, about 10^5 years or more, then the integration error starts to dominate and it is better to switch to a 3-step interpolation, because this will save CPU-time and does not affect accuracy. We also found (not shown) that the 2-step interpolation which requires less CPU-time than the 3-step interpolation will give the same accuracy after even longer integration times.

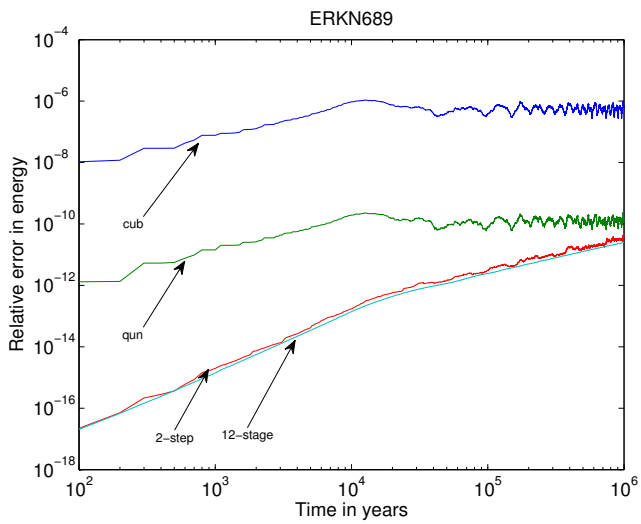

Fig. 6. The $H_{\rm rel}$ for *ERKN689* using cubic, quintic, 2-step Hermite interpolations, and the 12-stage interpolant over one million years of integration for the Jovian problem with local error tolerance of 10^{-10}

Figure 6 shows the relative error growth in energy for *ERKN689* using the cubic, quintic, and two-step Hermite interpolation schemes together with a 12-stage interpolant. For $TOL = 10^{-10}$, the average time-step is approximately 65 days. For the two-step interpolation and the 12-stage interpolant, there is no real difference

between these two graphs, unlike for *ERKN101217* using the 3-step interpolation and its interpolants as shown in Figure 5. The two errors increase side by side right from the beginning, indicating that there is no sense in using the 12-stage interpolant, because it costs more CPU-time than the 2-step interpolation.

We extend the above experiment for *ERKN689* by performing an integration of 10 million years. After approximately 7 million years (not shown) it is beneficial to switch to quintic interpolation, because it takes less CPU-time. The quintic Hermite interpolation is not under consideration for the *ERKN101217* integrator, because the total error differs by three orders of magnitude in comparison with *ERKN689*.

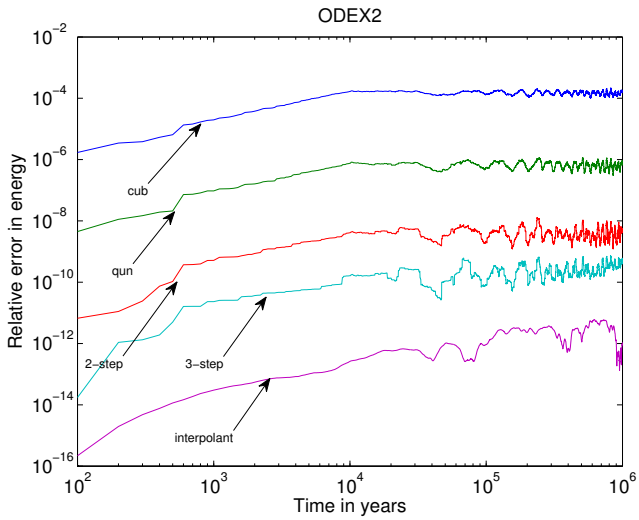

Fig. 7. The $H_{\rm rel}$ for *ODEX2* using cubic, quintic, 2-step, 3-step Hermite interpolations, and its built-in interpolant over one million years of integration for the Jovian problem with the local error tolerance of 10^{-16}

Figure 7 illustrates the relative error growth in energy for *ODEX2* using cubic, quintic, two-step and three-step interpolations together with its interpolant. On average, *ODEX2* uses a time-step of approximately 260 days if we set $TOL = 10^{-16}$; normally it is not recommended to take such a small tolerance, because it is very close to the machine precision and the round-off error could affect the results greatly. We observe that the graph for the interpolant shows oscillations, which indicates that the round-off error is significant. We also found that there are quite a few rejected time-steps. Combined with the asence of any rejected time-steps with either of the tolerances 10^{-14} and 10^{-15}, this is further evidence that the results for $TOL = 10^{-16}$ are affected by the round-off error. Nevertheless, we are gaining an accuracy of approximately 10^{-11} (determined by a linear least square fit using data obtained by continuous approximation). This accuracy for $TOL = 10^{-16}$ is approximately one order of magnitude better than the accuracy obtained with $TOL = 10^{-15}$, which is expected when reducing the tolerance by a factor of 10. This observation suggests that the round-off error is insignificant. The behaviour of the interpolant at the end of the integration shows an abrupt dip followed by rapid oscillations. We investigated this further by performing the experiments with the same tolerance for up to 10^8 years. While the oscillations continue, both $H_{\rm rel}$ and $L_{\rm rel}$ increase with the passage of time and the curve ends at a total error around 10^{-9} and 10^{-10}, respectively.

We performed the same experiments using the $\bar{S}$-*13* integrator with a step-size of four days, along with the cubic and quintic Hermite interpolation schemes. We found for the cubic interpolation that the total error in energy and angular momentum does not grow with t, but the maximum global error has a reasonably good agreement with the results obtained using a step-size sequence of five days in the experiments (see Figure 4). For the quintic Hermite interpolation, the error growth in energy and angular momentum is below linear.

Integrator	**Interpolation**	E_r	E_v	$H_{\rm rel}$	$L_{\rm rel}$
ERKN689	—	1.980	1.987	0.986	0.971
ERKN689	2-step	—	—	0.991	0.979
ERKN689	12-stage	—	—	0.985	0.970
ERKN1012	—	1.997	1.996	0.998	1.005
ERKN101217	3-step	—	—	0.677	0.683
ERKN101217	23-stage	—	—	0.998	1.005
ERKN101217	26-stage	—	—	0.997	1.004
ERKN101217	29-stage	—	—	0.997	1.005
ODEX2	—	1.514	1.457	0.327	0.474
ODEX2	3-step	—	—	0.589	0.632
ODEX2	interpolant	—	—	0.325	0.465
$\bar{S}$-*13*	—	1.377	1.542	0.689	0.718
$\bar{S}$-*13*	quintic	—	—	0.687	0.724

Table 2. The exponent b of the power law for the global error in position and velocity, and $H_{\rm rel}$ and $L_{\rm rel}$ over one million years for the Jovian problem. We used local error tolerances 10^{-11}, 10^{-10}, and 10^{-16} with different interpolation schemes, for the integrators *ERKN689*, *ERKN101217*, *ODEX2*, and $\bar{S}$-*13*, respectively. The step-size for $\bar{S}$-*13* is four days

Table 2 shows the exponent b in a linear least squares fit for the power law at^b. In particular, when low-order interpolation schemes are used with high-order variable-step-size integrators, the exponent b of the power law should be very close to zero. For example, for *ERKN689*, using the cubic and quintic Hermite interpolation schemes, the exponent b varies from 0.029 to 0.35 for $H_{\rm rel}$ and $L_{\rm rel}$. We already noted that the *ODEX2* integrator is affected by round-off error if $TOL = 10^{-16}$. To observe the extended behaviour of $H_{\rm rel}$ and $L_{\rm rel}$, we integrated *ODEX2* using $TOL = 10^{-16}$ over 10^7 years and observed the values for $H_{\rm rel}$ and $L_{\rm rel}$ as 0.76 and 0.89, respectively. We also repeated experiments for *ODEX2* with $TOL = 10^{-15}$ and 10^{-14} and observed that the value of the exponent increases. In particular, with $TOL = 10^{-14}$, the value of b is approximately 1.92 for the global error in position; for $H_{\rm rel}$ and $L_{\rm rel}$, it is approximately 0.99 and 1.01, respectively.

V. SUMMARY AND CONCLUSIONS

The main objective of this paper was to investigate the possibilities of replacing the interpolants of certain integrators by other interpolation schemes, perhaps at a cost of a little bit of accuracy. We analysed and compared the error growth for different combinations of numerical integrators and interpolation schemes. Our numerical testing involved comparing combinations of different integrators and interpolation schemes over a short time interval and several long time intervals, as long as 100 million years for the Jovian problem with local error tolerances ranging from 10^{-16} to 10^{-08}.

The interpolation schemes play a vital role in our work. The low-order interpolation schemes are unlikely to be used in practice. We have included them in our testing because we wanted to assess the effectiveness of these interpolation schemes. Especially, it makes sense to use a low-order interpolation scheme with a high-order Störmer method when the step-size is four days (for the Jovian problem), because this choice means the interpolation error is below machine precision. We performed many experiments using different interpolation schemes with different integrators both over a short time interval and several long time intervals of duration as long as 100 million years. For long term simulations, experiments were performed for the Jovian problem integrated up to 10^8 years. One of the strongest conclusions from these experiments is that the order of the continuous approximation should, with one notable exception in this paper, be compatible with the order of the integrator. The notable exception occurs for the higher order Störmer methods when used with small step-sizes. The order of the continuous approximation can be significantly lower than the order of the Störmer method. The reason why we get away with the low-order continuous approximation is because the $\bar{S}$*-13* integrator uses an artificially small step-size of four days to eliminate the truncation error at the machine precision and we are just left with the round-off error. However, if we want the truncation error to be 10^{-16}, then we might use the step-size of sixteen days rather than four days. When we use sixteen days, then the low-order continuous approximation (quintic Hermite interpolation) would fail.

To observe the behaviour of conserved quantities like energy and angular momentum using different interpolation schemes with different integrators, quite a few experiments were performed. All these experiments were performed for the Jovian problem over upto 10^8 years. One of the strong observations about all these experiments is that if a low order interpolation scheme is used with a high order integrator then the interpolation error for a particular time will dominate the integration error and will not increase with t. This is because the interpolation error doesn't grow with time. We investigated the relative error in energy for *ERKN101217* when used with cubic, quintic, two-step and three-step interpolation polynomials and the 3 interpolants, for *ERKN689* using the cubic, quintic, and two-step Hermite interpolation schemes together with a 12-stage interpolant, for *ODEX2* using cubic, quintic, two-step and three-step interpolations together with the interpolant, and for $\bar{S}$*-13* using cubic, and quintic interpolation schemes.

ERKN101217 across the integration of 10^6 years using $TOL = 10^{-11}$ on average takes approximately 290 days of time-step. We observed that the total error for all the low order interpolation schemes doesn't increase with the passage of time. With all three interpolants, we obtained the accuracy of approximately 10^{-11} and then with the 3-step interpolation scheme which is cheaper than all the three interpolants, the accuracy is about the same at the end of the integration, though initially it was dominated by the an interpolation error. So, about 10^4 years out of all these possibilities, it is better to use interpolant if the maximum accuracy is required but if it is integrated far enough, the integration error starts dominating, and then it is better to switch to a 3-step interpolation because that is going to save the CPU time and doesn't cost the accuracy. And if integrated even further, a 2-step interpolation is going to give the same accuracy, which is a little cheaper, than a 3-step interpolation.

ERKN689 with $TOL = 10^{-10}$ uses an average time-step of approximately 65 days. We observed that for a two-step interpolation and a 12-stage interpolant, there is not any real difference between the plotted graphs for the relative error in energy, unlike *ERKN101217* using a 3-step interpolation and interpolants, the two graphs go side by side right from the beginning, indicating right from the beginning there is no sense of using a 12-stage interpolant because it takes more CPU time in comparison with a 2-step interpolation. And after approximately 7 million years it is handy to switch to a quintic interpolation (which was not even under consideration with *ERKN101217* as these two integrators with the quintic Hermite interpolation are different total error by three orders of magnitude) as it costs less CPU time.

On average, *ODEX2* with $TOL = 10^{-16}$ uses a time-step of approximately 260 days. We observed that the graph for the interpolant shows oscillations, which is indicating the possible effect of the round-off error. We also observed that there are quite a few rejected time-steps. This combined with the absence of any time-steps with either of the tolerances 10^{-14} and 10^{-15} is further evidence that the results for $TOL = 10^{-16}$ are being affected by the round-off error. Although, it is affected by the round off error, yet we are gaining accuracy up to 10^{-11}, which is approximately one order of magnitude better than the accuracy obtained with $TOL = 10^{-15}$ and that is what you can expect by reducing the tolerance 10 times. So, reducing the tolerance from 10^{-15} to 10^{-16} is not making things difficult but we are gaining accuracy, which indicates that the round-off error is insignificant. For the interpolant, the behaviour of the plot at the end of the integration shows that the curve with oscillations is dipping-down. To observe the extended behavior we performed the experiment with the same tolerance for up to 10^8 years and it was observed that (with the oscillation going up and down) both H_{rel} and

$L_{\rm rel}$ increase with the passage of time and end-up with the total error around 10^{-9} and 10^{-10}, respectively. For $\bar{S}$-*13*, the integration was performed in double precision and we observed that for cubic interpolation the total error in energy and angular momentum doesn't grow with t but the achieved accuracy has a reasonably good agreement with the results obtained using 5 days in the experiment shown in Figure 4. For the quintic Hermite interpolation, the error growth in energy and angular momentum has been observed below the linear (the exponent of power law is less than 1 as mentioned in Table 2) error growth.

Acknowledgment. We are grateful to reviewer for their valuable suggestions to improve the quality of the article.

[1] T. S. Baker, J. R. Dormand, P. J. Prince, App. Num. Math. **29**, 171 (1999); https://doi.org/10.1016/S0168-9274(98)00065-8.

[2] J. Dormand, M. E. A. El-Mikkawy, P. Prince, J. Numer. Anal. **7**, 423 (1987); https://doi.org/10.1093/imanum/7.4.423.

[3] J. R. Dormand, P. J. Prince, Celest. Mech. **18**, 223 (1978); https://doi.org/10.1007/BF01230162.

[4] J. R. Dormand, P. J. Prince, Comp. Math. Appl. **12A**, 1007 (1986); https://doi.org/10.1016/0898-1221(86)90025-8.

[5] J. R. Dormand, P. J. Prince, Comp. Math. Appl. **13**, 937 (1987); https://doi.org/10.1016/0898-1221(87)90066-6.

[6] W. H. Enright, K. R. Jackson, S. P. Norsett, P. G. Thomsen, ACM Trans. Math. Softw. **12**, 193 (1986); https://doi.org/10.1145/7921.7923.

[7] K. R. Grazier, W. I. Newman, P. W. Sharp, Astron. J. **145**, 112 (2013): https://doi.org/10.1088/0004-6256/145/4/112.

[8] K. R. Grazier, W. I. Newman, J. M. Hyman, P. W. Sharp, ANZIAM J. **46**, C1086 (2005); https://doi.org/10.21914/anziamj.v46i0.1008.

[9] K. R. Grazier, W. I. Newman, W. M. Kaula, J. M. Hyman, Icarus **140**, 341 (1999); https://doi.org/10.1006/icar.1999.6146.

[10] E. Hairer, S. P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems* (Springer-Verlag, 1987).

[11] M. K. Horn, SIAM J. Num. Anal. **20**, 558 (1983); https://doi.org/10.1137/0720036.

[12] E. J. Nyström, Acta Soc. Sci. Fennicae **50**, 1 (1925); https://lib.ugent.be/catalog/rug01:001785116.

[13] L. F. Shampine, SIAM J. Num. Anal. **22**,1014 (1985); https://doi.org/10.1137/0722060.

[14] C. Störmer, Radium (Paris) **9**, 395 (1912); https://doi.org/10.1051/radium:01912009011039501.

[15] Ch. Tsitouras, G. Papageorgiou, Computing **43**, 255 (1990); https://doi.org/10.1007/BF02242920.

[16] J. H. Verner, SIAM J. Numer. Anal. **30**, 1446 (1990); https://doi.org/10.1137/0730075.

[17] S. Rehman, J. Comput. Appl. Math. **4**, 446 (2014); https://doi.org/10.4236/ajcm.2014.45037.

[18] K. R. Grazier, W. I. Newman, F. Varadi, W. M. Kaula, J. M. Hyman, Icarus **140**, 353 (1999); https://doi.org/10.1006/icar.1999.6146.

[19] D. R. Kirsh, M. Duncan, R. Brasser, H. F. Levison, Icarus **199**, 197 (2009); https://doi.org/10.1016/j.icarus.2008.05.028.

[20] M. S. Tiscareno, R. Malhotra, Astron. J. **138**, 827 (2009); https://doi.org/10.1088/0004-6256/138/3/827.

[21] P. S. Lykawka, J. Horner, B. W. Jones, T. Mukai, Month. Not. R. Ast. Soc. **4**, 1715 (2009); https://doi.org/10.1111/j.1365-2966.2009.15243.x.

[22] D. A. Minton, R. Malhotra, Icarus **207**, 744 (2010); https://doi.org/10.1016/j.icarus.2009.12.008.

[23] P. W. Sharp, ACM Trans. Math. Softw. **32**, 375 (2006); https://doi.org/10.1145/1163641.1163642.

НЕПЕРЕРВНІ НАБЛИЖЕННЯ ДЛЯ ДОВГОСТРОКОВИХ ЧИСЛОВИХ СИМУЛЯЦІЙ СОНЯЧНОЇ СИСТЕМИ

Ш. Рехман

Кафедра математики Інженерно-технологічного університету, Лахор, Пакистан

писано та проаналізовано ефективність різних комбінацій чотирьох числових інтеґраторів вищого порядку та до дев'яти схем інтерполяції, застосованих до задачі, яка включає Сонце й чотири газові ґіґанти (зовнішні планети), а саме: Юпітер, Сатурн, Уран та Нептун. Розглянуто схеми інтерполяції Ерміта з одним, двома та трьома часовими кроками й інтерполянтами для інтеґраторів ODEX2 та ERKN. Інтерполянти є особливим прикладом схеми інтерполяції, що дають наближення, неперервне протягом одного кроку й на всьому інтервалі інтеґрування. Інтерполянти досить вартісні порівняно з іншими схемами інтерполяції. Тому однією з цілей цієї праці є дослідження можливостей заміни інтерполянтів певних інтеґраторів іншими схемами інтерполяції, можливо, ціною незначної втрати точності. Експерименти проведено, щоб дослідити зростання похибок у положеннях, швидкостях і відносній похибці енерґії та кутового моменту з використанням різних комбінацій інтеґраторів і схем інтерполяції на великому інтервалі інтеґрування, аж до 100 мільйонів років для задачі Юпітера з локальним допуском помилок від 10^{-16} до 10^{-08}.

Ключові слова: проблема Юпітера, схеми інтерполяції, довгострокове моделювання.