Analysis for parareal algorithms applied to Hamiltonian differential equations

Martin J. Gander[a], Ernst Hairer[a]

[a]*Section de Mathématiques, Université de Genève, 2-4 rue du Lièvre, CH-1211 Genève 4, Switzerland. (Martin.Gander@unige.ch, Ernst.Hairer@unige.ch).*

Abstract

Long-time integrations are an important issue in the numerical solution of Hamiltonian systems. They are time consuming and it is natural to consider the use of parallel architectures for reasons of efficiency. In this context the parareal algorithm has been proposed by several authors.

The present work is a theoretical study of the parareal algorithm when it is applied to Hamiltonian differential equations. The idea of backward error analysis is employed to get insight into the long-time behaviour of numerical approximations. One of the main results is that convergence of the parareal iterations restricts the length of the time window. For nearly integrable systems its length is bounded by the square root of the inverse of the accuracy of the coarse integrator. The theoretical bounds are confirmed by numerical experiments.

Keywords: Hamiltonian differential equations, parareal algorithm, parallel architectures, symplectic methods, backward error analysis, long-time integration
2000 MSC: 65L05, 65Y05, 65P10

1. Introduction

Long-time integrations of differential equations can be time consuming and it is a natural idea to consider the use of computations in parallel. In this work we restrict our attention to Hamiltonian systems of ordinary differential equations

$$\dot{y} = f(y), \qquad f(y) = J^{-1}\nabla H(y), \qquad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{1}$$

where the vector $y = (p, q)$ collects momenta and positions, and the smooth scalar function $H(y)$ is called Hamiltonian or energy of the system. The symbol I in the structure matrix J is the d-dimensional identity matrix, and d is the number of degrees of freedom. The exact flow $\varphi_t(y)$ is for every t a symplectic transformation, which means that its derivative with respect to the initial value satisfies

$$\varphi_t'(y)^{\mathsf{T}} J \varphi_t'(y) = J. \tag{2}$$

It is well known that for long-time integrations the use of symplectic methods is recommended. This means that the discrete flow $y_{n+1} = \Phi_{\Delta T}(y_n)$ should share the property (2) with the exact flow. This then implies that the energy is nearly preserved along the numerical solution, see [7].

The parareal algorithm has been introduced by Lions, Maday, and Turinici in [11] as an approach for exploiting parallel architectures in the numerical solution of real time problems.

Among earlier attempts for parallel in time discretizations let us mention the work of Bellen and Zennaro [3], where a Steffensen-like iterative method is proposed for general difference equations, the parallel shooting technique of Chartier and Philippe [4], and the review article [9] on CWI (Amsterdam) contributions on parallel Runge–Kutta methods.

The parareal algorithm, originally proposed for the numerical treatment of partial differential equations, has also found applications in ordinary differential equations. During the last few years much research has been devoted to the long-time integration of Hamiltonian systems. The "symplectic parareal" of [2] is a non-iterative algorithm whose practical relevance has to be proven. Multi-time step parareal algorithms [1] are proposed for parallelizing in time molecular dynamics problems. The article [10] proposes a time-parallel algorithm for almost integrable Hamiltonian systems and has in mind very accurate computations in planetary motion. Our research is to a large extent motivated by [5], where symmetric parareal algorithms including projections are proposed for Hamiltonian systems.

After recalling the definition of the parareal algorithm for initial value problems of ordinary differential equations (Section 2), we present the main results, which are long-time error estimates for the parareal iterates for Hamiltonian systems, and numerical illustrations in Section 3. If ε denotes the accuracy of the coarse integrator, we show that convergence can be achieved on intervals whose length is restricted by $O(\varepsilon^{-1})$ for the harmonical oscillator, by $O(\varepsilon^{-1/2})$ for integrable systems, and by $O(1)$ for systems with chaotic solutions. The proofs are given in Section 4 for linear problems and in Sections 5 and 6 for nonlinear problems. Section 7 shows how the analysis extends to the symmetric parareal algorithm. For high accuracy computations we present a variant of the parareal algorithm in Section 8, which permits to use quadruple precision for the fine integrator and double precision for the coarse integrator.

2. Parareal algorithm for ordinary differential equations

Consider an initial value problem of ordinary differential equations $\dot{y} = f(y)$, $y(0) = y_0$. For a given step size ΔT we consider two discrete flow maps: a cheap approximation with low accuracy which we denote by $\varphi^G_{\Delta T}(y)$ (coarse integrator), and an accurate approximation which we denote by $\varphi^F_{\Delta T}(y)$ (fine integrator). The parareal algorithm, which is given by $u^k_0 = y_0$ for all k, and

$$\begin{aligned} u^0_{n+1} &= \varphi^G_{\Delta T}(u^0_n) \\ u^k_{n+1} &= \varphi^F_{\Delta T}(u^{k-1}_n) - \varphi^G_{\Delta T}(u^{k-1}_n) + \varphi^G_{\Delta T}(u^k_n), \end{aligned} \tag{3}$$

is expected to yield approximations that converge rapidly to the solution of the fine integrator. We use the notation

- c_G cost (cpu time) of one step of the coarse integrator $\varphi^G_{\Delta T}$,
- c_F cost (cpu time) of one step of the fine integrator $\varphi^F_{\Delta T}$,
- K maximal number of parareal iterations,
- N number of steps in one time window of length T, i.e., $N\Delta T = T$.

A *sequential computation* of the solution on an interval of length $T = N\Delta T$ with the accurate integrator $\varphi^F_{\Delta T}$ requires a total time

$$Nc_F \qquad \text{with} \qquad 1 \text{ processor.}$$

For the *standard application of the parareal algorithm* on the same interval one first computes sequentially the values u^0_n for all $n = 1, \ldots, N$. In the kth iteration, one computes all values

$\varphi^F_{\Delta T}(u^{k-1}_n)$ in parallel and subsequently the values u^k_n in a sequential manner. This requires a total time

$$Nc_G + K(Nc_G + c_F) \qquad \text{with} \qquad N \text{ processors.}$$

The performance can be improved, if the different processors are able to perform different tasks at the same time. The idea is to start the computation with the integrators $\varphi^F_{\Delta T}$ and $\varphi^G_{\Delta T}$ as soon as the initial values are available. Such a procedure is called *pipelined parareal algorithm* in [12] and requires a total time $Nc_G + K(c_G + c_F)$ with N processors. A minor improvement is possible to perform the computation in a total time

$$Nc_G + Kc_F \qquad \text{with} \qquad (q+1)K + 1 \text{ processors,}$$

where q denotes the smallest positive integer satisfying $c_F \le qc_G$. This can be achieved as follows: on the first processor one computes sequentially the values u^0_{n+1} for $n = 0, \ldots, N-1$; in the kth iteration, q processors are used to compute $\varphi^F_{\Delta T}(u^{k-1}_n)$ for $n = k-1, \ldots, N-1$, starting the computation as soon as u^{k-1}_n is available, and a further processor computes sequentially the values $\varphi^G_{\Delta T}(u^k_n)$ and u^k_{n+1} for $n = k, \ldots, N-1$. The first processor requires the time Nc_G and every parareal iteration requires the additional time c_F.

In general it is not possible to treat the whole interval $[0, T_{\text{end}}]$, where the problem has to be solved, in one piece. One has to divide it into time windows on which the parareal algorithm can be applied. To get a significant benefit from the use of parallel architectures, the length T of the time window should be as large as possible. It is one of the main results of our research that this length T cannot be arbitrarily large for Hamiltonian systems and is restricted in terms of the accuracy of the coarse integrator.

3. Main results illustrated by numerical experiments

The parareal iterates approximate the solution that corresponds to a sequential application of the fine integrator. We denote this solution by $y_F(t_n) = y_n$ for $t_n = n\Delta T$, i.e., it is given by $y_{n+1} = \varphi^F_{\Delta T}(y_n)$. Throughout this article we assume that $\varphi^F_{\Delta T}$ is a symplectic method of order $\ge r$. This implies that for the problem (1) there exist functions $H_j(y)$, such that the truncated series

$$H_F(y) = H(y) + \Delta T^r H_r(y) + \Delta T^{r+1} H_{r+1}(y) + \ldots + \Delta T^M H_M(y), \tag{4}$$

called modified Hamiltonian, satisfies

$$H_F(y_F(t_n)) = H_F(y_0) + O(t_n \Delta T^M).$$

The constant symbolized by $O(\cdot)$ depends on M, which can be choosen arbitrarily, but it is independent of ΔT and n, see [7, Chapter IX]. As a consequence we have near energy preservation

$$H(y_F(t_n)) = H(y_0) + O(\Delta T^r) + O(t_n \Delta T^M)$$

on intervals where the second error term can be neglected with respect to the first one. Notice that for symmetric methods the expansion (4) is in even powers of ΔT.

In the following we consider several typical Hamiltonian systems, we present sharp estimates of the error in the solution and in the energy on intervals as long as possible, and we verify the results by numerical experiments. We start with the harmonic oscillator which is a Hamiltonian system with

$$H(p, q) = \frac{1}{2}(p^2 + q^2), \tag{5}$$

and which results in a linear differential equation with constant coefficients.

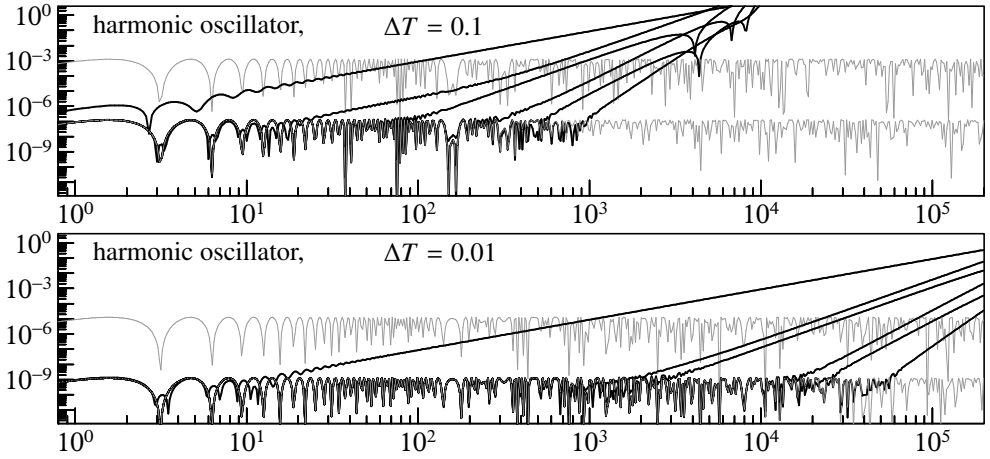

Figure 1: Error in the Hamiltonian for the parareal algorithm applied to the harmonic oscillator.

Theorem 1 (harmonic oscillator). *Consider a coarse integrator of order r, and assume that the fine integrator is symplectic for (5) and of order ≥ r. For the harmonic oscillator the numerical solution $u_n^k = (p_n^k, q_n^k)$ of the k-th parareal iteration satisfies, for $t = t_n = n\Delta T$ and for sufficiently small ΔT,*

$$\|u_n^k - y_F(t)\| \le C\,\frac{(c\,t\,\Delta T^r)^{k+1}}{(k+1)!} \tag{6}$$

as long as $t\,\Delta T^r$ is bounded. The constants are independent of n, k, and ΔT. The error in the energy, $\left|H(p_n^k, q_n^k) - H(p_0, q_0)\right|$, admits the same bound with a possibly different constant C.

If both integrators are analytic (which is true for Runge–Kutta methods) and if r is even, the error in the energy is bounded by

$$\left|H(p_n^k, q_n^k) - H(p_0, q_0)\right| \le \begin{cases} C\,\dfrac{(c\,t\,\Delta T^r)^{k+1}}{(k+1)!} & \text{if } k \text{ is odd,} \\[2ex] C\,\dfrac{(c\,t\,\Delta T^r)^{k+1}}{(k+1)!}\left(\Delta T + \dfrac{c\,t\,\Delta T^r}{k+2}\right) & \text{if } k \text{ is even.} \end{cases} \tag{7}$$

Notice that the coarse integrator need not be symplectic to get the estimates of Theorem 1, which are seen to be sharp in the following example. We apply the parareal algorithm with the Störmer–Verlet method (see [7]) as basic method: one step with step size ΔT for $\varphi_{\Delta T}^G$, and 100 steps with step size $\Delta T/100$ for $\varphi_{\Delta T}^F$. The errors of the energy along the numerical solution are shown for two different ΔT, so that the dependence on ΔT can be better observed. In all figures the numerical energy for $\varphi_{\Delta T}^G$ and $\varphi_{\Delta T}^F$ is drawn in gray, that for the parareal iterations in black.

Figure 1 shows the numerical Hamiltonian corresponding to the harmonic oscillator with initial values $q(0) = 1$, $p(0) = 0$. Although the basic integrators $\varphi_{\Delta T}^G$ and $\varphi_{\Delta T}^F$ are symplectic and show a bounded error in the Hamiltonian (without any drift), the parareal iterations have a drift. The error is seen to behave linearly (in double logarithmic scale) with slopes $2, 4, 4, 6, 6, 8$ for the k-th iterates ($k = 1, \ldots, 6$). This behavior can be observed on intervals of length t, where $c\,t\,\Delta T^2 \le 1$. Indeed, when dividing the step size ΔT by 10, we observe a comparable size of the error at a time instant that is 100 times larger. This perfectly agrees with the theoretical estimates of Theorem 1. Note that $r = 2$ is the order of the Störmer–Verlet scheme.

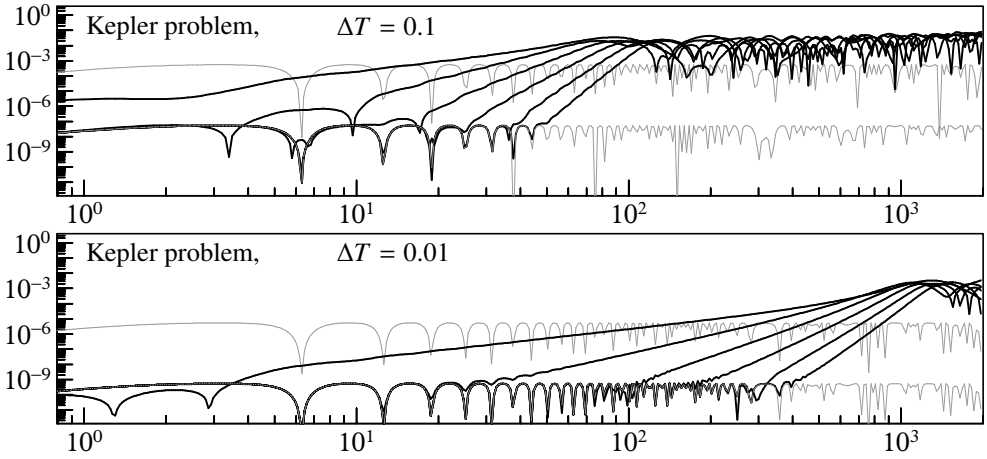

Figure 2: Error in the Hamiltonian for the parareal algorithm applied to the Kepler problem.

We next consider completely integrable Hamiltonian systems. For such systems there exists a symplectic change of coordinates to action-angle variables. The action variables are conserved quantities and the angle variables are linear functions of time. This implies that perturbations in the initial values propagate at most linearly with time [7, Chapter X].

Theorem 2 (completely integrable systems). *Consider a coarse integrator of order r, a fine integrator of order $\geq r$, and assume that both of them are symplectic. For a completely integrable Hamiltonian system the numerical solution $u_n^k = (p_n^k, q_n^k)$ of the k-th parareal iteration satisfies, for $t = t_n = n\Delta T$ and for sufficiently small ΔT,*

$$\|u_n^k - y_F(t)\| \leq c_k\, t\, \Delta T^r\, (t^2\, \Delta T^r)^k, \tag{8}$$

where c_k depends on the problem, but is independent of n and ΔT as long as $t^2\, \Delta T^r$ is bounded and $t \geq 1$.

For the error in the energy we have the estimate

$$\left|H_F(p_n^k, q_n^k) - H_F(p_0, q_0)\right| \leq c_k\, \Delta T^r\, (t^2\, \Delta T^r)^k \tag{9}$$

with a possibly different constant c_k.

In the estimates of Theorems 2 and 3 we do not investigate further the dependence of the constants c_k on k. This would require the study of the constants C_k in Corollary 5 and assumptions on the derivatives of the vector field $f(y)$. We believe, and this is confirmed by the discussion of Section 8, that the parareal algorithm is efficient only if the desired accuracy is achieved after a few iterations. In this case the dependence on k of the constants c_k is not relevant.

A typical example of a completely integrable system is the Kepler problem with

$$H(p, q) = \frac{1}{2}(p_1^2 + p_2^2) - \frac{1}{\sqrt{q_1^2 + q_2^2}}. \tag{10}$$

The numerical result is shown in Figure 2 for initial values $q_1(0) = 1 - e$, $q_2(0) = 0$, $p_1(0) = 0$, and $p_2(0) = \sqrt{(1+e)/(1-e)}$, so that the solution is an ellipse with eccentricity $e = 0.1$.

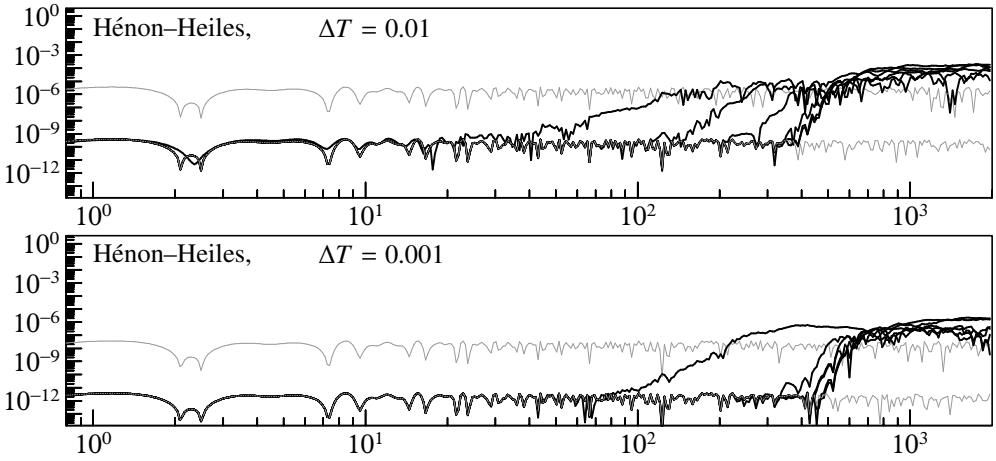

Figure 3: Error in the Hamiltonian for the parareal algorithm applied to the Hénon–Heiles problem.

The long-time behavior is significantly different from that for the harmonic oscillator. We again notice that on intervals, where the solution is reasonably accurate, the error behaves linearly in double logarithmic scale, but this time the slopes are $2, 4, 6, 8, 10, 12$. When dividing the step size ΔT by 10, we observe a comparable size of the error at a time instant that is only 10 times larger. This is in complete agreement with the estimates of Theorem 2 for $r = 2$.

We finally consider Hamiltonian systems with chaotic solutions. The total energy is preserved along the solution, but perturbations in initial values are propagated exponentially fast like e^{Lt}. We assume the Lyapunov exponent L to be of moderate size, and that the solution does not have high oscillations.

Theorem 3 (general Hamiltonian systems). *Consider a coarse integrator of order r, a fine integrator of order $\geq r$, and assume that both of them are symplectic. For a Hamiltonian system, where L is a Lipschitz constant of the vector field, the numerical solution $u_n^k = (p_n^k, q_n^k)$ of the k-th parareal iteration satisfies, for $t = t_n = n\Delta T$ and for sufficiently small ΔT,*

$$\|u_n^k - y_F(t)\| \leq c_k \, \mathrm{e}^{Lt} \, (t \, \Delta T^r)^{k+1}, \tag{11}$$

where c_k is independent of n and ΔT as long as $t \, \Delta T^r$ is bounded.

For the error in the energy we have the estimate

$$\left| H_F(p_n^k, q_n^k) - H_F(p_0, q_0) \right| \leq c_k \, \Delta T^r \, \mathrm{e}^{Lt} \, (t \, \Delta T^r)^k. \tag{12}$$

The estimate (11) of the global error is the same as that for general differential equations given in [6]. There, it is shown that $c_k = Cc^{k+1}/(k+1)!$ with k-independent constants c and C.

A problem with a chaotic solution is the Hénon–Heiles equation given by

$$H(p,q) = \frac{1}{2}(p_1^2 + p_2^2) + U(q_1, q_2), \qquad U(q_1, q_2) = \frac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3} q_2^3 \tag{13}$$

with initial values $q_1(0) = 0$, $q_2(0) = 0.2$, $p_2(0) = 0.2$, and the positive value of $p_1(0)$ is chosen such that $H(p(0), q(0)) = 1/8$, see [7, Sect. I.3]. The numerical Hamiltonian of the parareal iterates are shown in Figure 3. In this situation, the error behavior is less regular and one cannot

easily discover slopes depending only on the iteration number and not on the step size. We have considered smaller step sizes than before to demonstrate that the interval, where the parareal algorithm improves the solution, does not increase for smaller ΔT. This is due to the factor e^{Lt} in the estimates of Theorem 3.

4. Analysis for the harmonic oscillator

The differential equation for the harmonic oscillator is linear, and every reasonable numerical integrator will depend linearly on the initial values. We therefore assume that $\varphi^G_{\Delta T}(y) = G(\Delta T)y$ and $\varphi^F_{\Delta T}(y) = F(\Delta T)y$ with 2×2 matrices $G(\Delta T)$ and $F(\Delta T)$.

The parareal algorithm (3) yields the recursion

$$u^0_{n+1} = G(\Delta T)\, u^0_n, \qquad u^k_{n+1} = G(\Delta T)\, u^k_n + \big(F(\Delta T) - G(\Delta T)\big)\, u^{k-1}_n$$

with $u^k_0 = y_0$ for all k. With the generating functions $u_n(\zeta) = \sum_{k\ge 0} u^k_n\, \zeta^k$ the parareal recursion can be written as

$$u_{n+1}(\zeta) = \Big(G(\Delta T) + \zeta\big(F(\Delta T) - G(\Delta T)\big)\Big)\, u_n(\zeta), \qquad u_0(\zeta) = \frac{1}{1-\zeta}\, y_0,$$

which gives

$$u_n(\zeta) = \Big(G(\Delta T) + \zeta\big(F(\Delta T) - G(\Delta T)\big)\Big)^n \frac{1}{1-\zeta}\, y_0.$$

The binomial theorem can be applied for obtaining explicit formulas for the numerical approximations u^k_n. Assuming that the matrices $G(\Delta T)$ and $F(\Delta T)$ commute (this is the case for all analytic methods, such as Runge–Kutta methods) we get

$$\begin{aligned} u^k_n &= \sum_{j=0}^{k} \binom{n}{j} G(\Delta T)^{n-j} \Big(F(\Delta T) - G(\Delta T)\Big)^j y_0 \\ &= F(\Delta T)^n\, y_0 \;-\; \sum_{j=k+1}^{n} \binom{n}{j} G(\Delta T)^{n-j} \Big(F(\Delta T) - G(\Delta T)\Big)^j y_0. \end{aligned} \tag{14}$$

If the coarse integrator is of order r and the fine integrator of order $\ge r$, we have $G(\Delta T) = F(\Delta T) + \Delta T^{r+1} D(\Delta T)$ with bounded $D(\Delta T)$. As long as $n\,\Delta T^{r+1}$ is small enough, the leading terms in the last sum of (14) are those with small j. Moreover, we have $G(\Delta T)^n = F(\Delta T)^n (I + O(\Delta T)^{r+1})^n$, and symplecticity of the fine integrator implies that $F(\Delta T)^n$ is bounded uniformly in n and ΔT. Restricting the analysis to time intervals $t_n = n\Delta T \le t$, such that $t\,\Delta T^r$ is bounded, we thus obtain the estimate (6) of Theorem 1. Taking norms, this part of the proof is extended straightforwardly to the general case, where $G(\Delta T)$ and $F(\Delta T)$ do not commute.

To get the improved error estimates (7), we write the differential equation of the harmonic oscillator as $\dot y = \mathrm{i}y$ with $y = p + \mathrm{i}q$. We also consider the numerical approximations $y_n = p_n + \mathrm{i}q_n$ and u^k_n as complex numbers. We assume that the integrators are analytic, so that $G(\Delta T)$ and $F(\Delta T)$ are complex numbers of the form $1 + \mathrm{i}\Delta T + O(\Delta T^2)$, and $D(\Delta T) = \mathrm{i}^{r+1} d_0 + O(\Delta T)$ with real d_0. We then have

$$\frac{G(\Delta T)}{F(\Delta T)} = 1 + (\mathrm{i}\Delta T)^{r+1} + O(\Delta T^{r+2}), \qquad \left(\frac{G(\Delta T)}{F(\Delta T)}\right)^n = \alpha_n + \mathrm{i}\,\beta_n(t\,\Delta T^r)$$

with real α_n, β_n that are bounded uniformly in n as long as $t\,\Delta T^r$ is bounded. Using these estimates for the dominating term $j = k+1$ in (14) and then also for the remaining terms shows that for methods with even order r the approximation u_n^k satisfies

$$u_n^k = F(\Delta T)^n \Big(1 + \mathrm{i}^{k+1}\,\frac{(c\,t\,\Delta T^r)^{k+1}}{(k+1)!}\Big(e_{1n}^k + \Delta T\, e_{2n}^k\Big) + \mathrm{i}^{k+2}\,\frac{(c\,t\,\Delta T^r)^{k+2}}{(k+2)!}\,e_{3n}^k\Big) y_0$$

with real c, e_{1n}^k and complex e_{2n}^k, e_{3n}^k. We have $|c| \le 2|d_0|$ and, for sufficiently small ΔT, the expressions $e_{1n}^k, e_{2n}^k, e_{3n}^k$ are bounded independently of n and k as long as $t\,\Delta T^r$ is bounded. Since the accurate integrator is symplectic for linear problems, we have $|F(\Delta T)| = 1$, which proves the estimate (7) of Theorem 1.

5. Nonlinear Hamiltonian systems – backward error analysis

We assume that the accurate integrator $y_{n+1} = \varphi_{\Delta T}^F(y_n)$ can formally[1] be interpreted as the exact flow of a modified differential equation (see [7, Chapter IX])

$$\dot y = f_{\Delta T}(y), \qquad f_{\Delta T}(y) = f(y) + \Delta T^r f_r(y) + \Delta T^{r+1} f_{r+1}(y) + \dots\,, \tag{15}$$

where r denotes the order of the method. This means that formally $y_n = y_F(t_n)$ with $t_n = n\Delta T$, where $y_F(t)$ is the solution of (15). Furthermore, we assume that the numerical solution of the less accurate integrator $y_{n+1} = \varphi_{\Delta T}^G(y_n)$ can be formally considered as the exact flow of

$$\dot y = f_{\Delta T}(y) + \Delta T^r d_{\Delta T}(y), \qquad d_{\Delta T}(y) = d_r(y) + \Delta T\, d_{r+1}(y) + \Delta T^2 d_{r+2}(y) + \dots \tag{16}$$

Following the work of [5], we consider the parareal algorithm (3) as a one-step method

$$\Psi_{\Delta T} : (u_n^0, u_n^1, \dots, u_n^K) \mapsto (u_{n+1}^0, u_{n+1}^1, \dots, u_{n+1}^K) \tag{17}$$

on the vector space that consists of $K+1$ copies of the vector space where the original problem is defined. We next derive the modified differential equation of this discrete flow. Obviously, it is equal to (16) for the first component $\{u_n^0\}$.

For the difference between the coarse and accurate solutions we have by Taylor expansion

$$\varphi_{\Delta T}^G(y) - \varphi_{\Delta T}^F(y) = \Delta T^{r+1} e_r(y) + \Delta T^{r+2} e_{r+1}(y) + \dots, \tag{18}$$

where

$$\begin{aligned} e_r(y) &= d_r(y) \\ e_{r+1}(y) &= d_{r+1}(y) + \frac{1}{2}\Big(f'(y)d_r(y) + d_r'(y)f(y)\Big). \end{aligned}$$

Writing the parareal algorithm (3) as $u_{n+1}^k = \varphi_{\Delta T}^P(u_n^{k-1}, u_n^k)$ with $\varphi_{\Delta T}^P(u,v) = \varphi_{\Delta T}^F(u) - \varphi_{\Delta T}^G(u) + \varphi_{\Delta T}^G(v)$, its discrete flow becomes

$$\varphi_{\Delta T}^P(u^{k-1}, u^k) = \varphi_{\Delta T}^F(u^k) + \sum_{j\ge r} \Delta T^{j+1}\Big(e_j(u^k) - e_j(u^{k-1})\Big). \tag{19}$$

[1] With the term "formally" we mean that we have identity of the Taylor series without taking care about convergence. In fact, the series in the modified differential equation is an asymptotic expansion and diverges for nearly all nonlinear problems.

We are now ready to compute the modified differential equation for the parareal algorithm. This means that we search for differential equations with solutions $u^k(t)$, such that the numerical approximations of the parareal iterates formally satisfy $u^k(t_n) = u_n^k$. For this we make the ansatz

$$\dot{u}^k = f_{\Delta T}(u^k) + \Delta T^r g_r^k(u^{k-1}, u^k) + \Delta T^{r+1} g_{r+1}^k(u^{k-1}, u^k) + \dots, \tag{20}$$

we compute the Taylor expansion of its solution on an interval of length ΔT, and we identify it with the discrete flow (19) of the parareal algorithm. This yields, for $k \geq 1$,

$$\begin{aligned} g_r^k(u^{k-1}, u^k) &= e_r(u^k) - e_r(u^{k-1}) \\ g_{r+1}^k(u^{k-1}, u^k) &= e_{r+1}(u^k) - e_{r+1}(u^{k-1}) - \frac{1}{2}\Big(f'(u^k) g_r^k(u^{k-1}, u^k) \\ &\qquad + \frac{\partial g_r^k}{\partial u^{k-1}}(u^{k-1}, u^k) f(u^{k-1}) + \frac{\partial g_r^k}{\partial u^k}(u^{k-1}, u^k) f(u^k) \Big). \end{aligned}$$

We remark that for $j \geq 2r+1$ and $k \geq 2$ the functions g_j^k depend in addition to u^{k-1}, u^k also on u^{k-2}, for $j \geq 3r+2$ and $k \geq 3$ also on u^{k-3}, etc.

Symmetric methods and reversible systems. For symmetric methods the order r is even, and the modified differential equation is known to have an expansion in even powers of ΔT (see [7, Chapter IX]). Assuming that both methods $\varphi_{\Delta T}^F$ and $\varphi_{\Delta T}^G$ are symmetric, we have $f_{r+1}(y) = 0$ and $d_{r+1}(y) = 0$. This implies that

$$g_{r+1}^k(u^{k-1}, u^k) = \frac{1}{2}\Big(f'(u^k) - f'(u^{k-1}) \Big) d_r(u^{k-1}),$$

which does not vanish in general. Consequently, the modified differential equation (20) is not an expansion in even powers of ΔT. Therefore, we cannot expect any improved behavior for reversible systems, which are problems with symmetries in the vector field, see [7, Chapter V].

Hamiltonian systems and symplectic methods. Consider Hamiltonian vector fields, that is $f(y) = J^{-1}\nabla H(y)$. If both methods $\varphi_{\Delta T}^F$ and $\varphi_{\Delta T}^G$ are symplectic, then their modified differential equations are Hamiltonian (see [7, Chapter IX]). This means that, for $j \geq r$,

$$f_j(y) = J^{-1}\nabla H_j(y), \qquad d_j(y) = J^{-1}\nabla K_j(y)$$

with smooth scalar functions $H_j(y)$ and $K_j(y)$. Denoting the Hamiltonian of the vector field (15) by $H_F(y) = H(y) + \Delta T^r H_r(y) + \dots$, the modified differential equation of (17) becomes, for $K = 1$,

$$\begin{pmatrix} \dot{u}^0 \\ \dot{u}^1 \end{pmatrix} = \begin{pmatrix} J^{-1} & 0 \\ 0 & J^{-1} \end{pmatrix} \begin{pmatrix} \nabla H_F(u^0) \\ \nabla H_F(u^1) \end{pmatrix} + \Delta T^r \begin{pmatrix} J^{-1} & 0 \\ -J^{-1} & J^{-1} \end{pmatrix} \begin{pmatrix} \nabla K_r(u^0) \\ \nabla K_r(u^1) \end{pmatrix} + \dots . \tag{21}$$

The perturbation term is not Hamiltonian, so that the one-step mapping (17) cannot be a symplectic transformation.

5.1. Structure of the modified differential equation

We take a closer look at the modified differential equation (20) of the parareal algorithm. We already mentioned earlier that the coefficient functions not only depend on u^{k-1} and u^k, but those

with larger index also depend on u^j with $j \le k-2$. We introduce the notation $\mathbf{u}^k = (u^0, \dots, u^k)$, and we write the modified differential equation as

$$\dot{u}^k = f_{\Delta T}(u^k) + \Delta T^r g_r^k(\mathbf{u}^k) + \Delta T^{r+1} g_{r+1}^k(\mathbf{u}^k) + \dots \ . \tag{22}$$

The computation of the first perturbation terms shows that they are a linear combination of expressions appearing as pairs. We therefore introduce the following notation, for $l \le k$,

$$g \in D_k(u^{l-1}, u^l) \quad \Leftrightarrow \quad g(\mathbf{u}^k) = a(\mathbf{u}^k, u^l) - a(\mathbf{u}^k, u^{l-1})$$

with a smooth function a. We also let $D_k(u^{l-2}, u^{l-1}, u^l) = D_k(u^{l-2}, u^{l-1}) \cup D_k(u^{l-1}, u^l)$ and, in general, $D_k(u^{l-m}, \dots, u^l) = D_k(u^{l-m}, u^{l-m+1}) \cup D_k(u^{l-m+1}, \dots, u^l)$.

Lemma 4. *Let $g_j^k(\mathbf{u}^k)$ be the coefficient functions of the modified differential equation (22) of the parareal algorithm. Then g_j^k is a linear combination of elements in*

$$\begin{aligned} D_k(u^{k-1}, u^k) \qquad & \text{for } \ r \le j < 2r+1, \ k \ge 1 \\ D_k(u^{k-2}, u^{k-1}, u^k) \qquad & \text{for } \ 2r+1 \le j < 3r+2, \ k \ge 2, \\ D_k(u^{k-m}, \dots, u^k) \qquad & \text{for } \ m(r+1)-1 \le j < m(r+1)+r, \ k \ge m. \end{aligned}$$

Proof. We proceed by induction on j. For $j = r$ the statement is obvious, because $g_r^k(\mathbf{u}^k) = e_r(u^k) - e_r(u^{k-1})$. For $j > r$, the functions $g_j^k(\mathbf{u}^k)$ are obtained recursively by comparing like powers of ΔT in the expansion of the solution at $t + \Delta T$ of (22) and that of the discrete mapping (19). The coefficient functions of (19) are all in $D_k(u^{k-1}, u^k)$. It is therefore sufficient to prove that the expressions arising from the derivatives of the solution of (22) are of the desired form.

In the following we collect all perturbation terms of (22) in $\Delta T^r g_{\Delta T}^k(\mathbf{u}^k)$, and we denote by $a'(\mathbf{u}^k, u^l)$ the derivative with respect to the second argument. If $g(\mathbf{u}^k)$ is a function in $D_k(u^{l-1}, u^l)$, then its total derivative is

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} g(\mathbf{u}^k) &= a'(\mathbf{u}^k, u^l)\, \dot{u}^l - a'(\mathbf{u}^k, u^{l-1})\, \dot{u}^{l-1} + \dots \\ &= \Delta T^r \Big(a'(\mathbf{u}^k, u^l)\, g_{\Delta T}^l(\mathbf{u}^l) - a'(\mathbf{u}^k, u^{l-1})\, g_{\Delta T}^{l-1}(\mathbf{u}^{l-1}) \Big) + \dots \ , \end{aligned} \tag{23}$$

where three dots represent an expansion in powers of ΔT, whose coefficients are functions lying in $D_k(u^{l-1}, u^l)$.

The function $g_{r+1}^k(\mathbf{u}^k)$ is a linear combination of $e_{r+1}(u^k) - e_{r+1}(u^{k-1})$, and of expressions coming from $\frac{1}{2!}\Delta T^2\, \ddot{u}^k$, which are $f'(u^k) g_r^k(\mathbf{u}^k)$ and the ΔT-independent term of the total derivative of $g_r^k(\mathbf{u}^k)$. It follows from (23) that all these functions are in $D_k(u^{k-1}, u^k)$. The same argument is valid for $g_j^k(\mathbf{u}^k)$ as long as $j < 2r+1$. This proves the first statement of the lemma.

For $j = 2r+1$ and $k \ge 2$ the ΔT^r terms of (23) are involved for the first time. We have just seen that $g_j^k(\mathbf{u}^k)$ is in $D_k(u^{k-1}, u^k)$ for $j < 2r+1$. This implies that $a'(\mathbf{u}^k, u^k)\, g_{\Delta T}^k(\mathbf{u}^k)$ and $a'(\mathbf{u}^k, u^{k-1})\, g_{\Delta T}^{k-1}(\mathbf{u}^{k-1})$ are in $D_k(u^{k-1}, u^k)$ and $D_k(u^{k-2}, u^{k-1})$, respectively. The same argument can be applied as long as $j < 3r+2$, which proves the second statement of the lemma.

For $j = 3r+2$ and $k \ge 3$, an expression of the form $a'(\mathbf{u}^k, u^{k-2}) g_{\Delta T}^{k-2}(\mathbf{u}^{k-2})$ will be involved inducing the presence of functions in $D_k(u^{k-3}, u^{k-2})$. Continuing the induction argument in this way completes the proof of the lemma. □

For rigorous error estimates based on backward error analysis, the infinite series in the modified differential equation (22) has to be truncated. To get exponentially small error estimates, one has to consider J terms, where J is proportional to ΔT^{-1} (see [7, Section IX.7]). Since we are not interested in exponentially small error estimates, it is sufficient to consider a truncation index J that is independent of ΔT (see [7, Section X.3] for the case of integrable Hamiltonian systems). It has to be at least $(r+1)(K+1)$, where r is the order of the coarse integrator and K is the maximal number of parareal iterations.

Corollary 5. *For the parareal algorithm, the truncated perturbation*

$$\Delta T^r g^k_{\Delta T}(\mathbf{u}^k) = \Delta T^r g^k_r(\mathbf{u}^k) + \Delta T^{r+1} g^k_{r+1}(\mathbf{u}^k) + \ldots + \Delta T^J g^k_J(\mathbf{u}^k)$$

of the modified differential equation (22) admits for $k \geq 1$ the bound

$$\|g^k_{\Delta T}(\mathbf{u}^k)\| \leq C_k \left(\|u^k - u^{k-1}\| + \Delta T^{r+1} \|u^{k-1} - u^{k-2}\| + \ldots + \Delta T^{(k-1)(r+1)} \|u^1 - u^0\| \right)$$

as long as u^l, for $l = 1, \ldots, k$, remain in a compact set of the phase space.

Proof. The mean value theorem implies that functions of $D_k(u^{k-1}, u^k)$ are bounded by a constant times $\|u^k - u^{k-1}\|$, if the arguments are restricted to a compact set. By Lemma 4 we thus have $\|g^k_j(\mathbf{u}^k)\| \leq C \|u^k - u^{k-1}\|$ for $j = r, \ldots, 2r$, $\|g^k_j(\mathbf{u}^k)\| \leq C (\|u^k - u^{k-1}\| + \|u^{k-1} - u^{k-2}\|)$ for $j = 2r+1, \ldots, 3r+1$, and so on. This implies the statement of the corollary. □

6. Long-time error estimates

The analysis for the long-time behavior of the harmonic oscillator (Section 4) does not permit to draw conclusions for nonlinear problems, not even for integrable Hamiltonian systems like the Kepler problem. This can be seen from the different behavior of the errors in the Hamiltonian shown in Figures 1 and 2. In this section we prove the results announced in Section 3.

6.1. Estimates for the global error

We let $y_F(t) = y(t, t_0, y_0)$ be the solution of the modified differential equation (15) for the accurate integrator, and we consider successively the modified equations (20) for the parareal algorithm. Since $\Delta T^r g^k_{\Delta T}(\mathbf{u}^k(t))$ is the defect when the function $u^k(t)$ is inserted into (15), the nonlinear variation of constants formula (see for example [8, Section I.14]) yields

$$u^k(t) - y_F(t) \;=\; \Delta T^r \int_0^t \frac{\partial y}{\partial y_0}\Big(t, s, u^k(s)\Big)\, g^k_{\Delta T}(\mathbf{u}^k(s))\, \mathrm{d}s. \tag{24}$$

This is an implicit equation for $u^k(t)$ in fixed point form. Due to the factor ΔT^r we have a contraction for sufficiently small ΔT.

General situation. We let $f_{\Delta T}(y)$ satisfy a Lipschitz condition with constant L on a compact set. As long as all considered solution approximations remain in this compact set we have

$$\left\| \frac{\partial y}{\partial y_0}\Big(t, s, u^k(s)\Big) \right\| \leq \hat{c}\, \mathrm{e}^{L(t-s)}, \qquad \|u^0(t) - y_F(t)\| \leq c_0\, t\, \mathrm{e}^{Lt} \Delta T^r. \tag{25}$$

The second relation follows from standard convergence estimates (e.g., Theorem II.3.4 of [8]) and the use of $(\mathrm{e}^{Lt} - 1)/L \leq t\, \mathrm{e}^{Lt}$ for $t \geq 0$.

Our proof of (11) is by induction on k. For $k = 0$ the statement is precisely the second estimate of (25). Assume now that (11) holds up to $k-1$. By Corollary 5 and the triangle inequality we then have

$$\begin{aligned}\left\|g^k_{\Delta T}(\mathbf{u}^k(t))\right\| &\le C_k\Big(\|u^k(t)-y_F(t)\| + (1+\Delta T^{r+1})\,c_{k-1}\,t^k \mathrm{e}^{Lt}\Delta T^{rk} \\ &\qquad + (1+\Delta T^{r+1})\,c_{k-2}\,t^{k-1}\mathrm{e}^{Lt}\Delta T^{rk+1} + \ldots + c_0\,t\,\mathrm{e}^{Lt}\Delta T^{rk+k-1}\Big) \\ &\le C_k\Big(\|u^k(t)-y_F(t)\| + p(t)\,\mathrm{e}^{Lt}\Delta T^{rk}\Big) \qquad (26)\end{aligned}$$

with $p(t) = c_{k-1}\,(1+\Delta T^{r+1})\,t^k + c_{k-2}\,(1+\Delta T^{r+1})\,t^{k-1}\Delta T + \ldots + c_0\,t\,\Delta T^{k-1}$. For the function $m(t) = \mathrm{e}^{-Lt}\,\|u^k(t) - y_F(t)\|$, the inequalities (25) and (26) inserted into (24) yield

$$m(t) \le \alpha \int_0^t \Big(m(s) + p(s)\,\Delta T^{rk}\Big)\,\mathrm{d}s, \qquad \alpha = \hat{c}\,C_k\Delta T^r.$$

Applying Gronwall's inequality we obtain

$$m(t) \le \alpha \int_0^t \mathrm{e}^{\alpha(t-s)} p(s)\,\Delta T^{rk}\,\mathrm{d}s.$$

For bounded $t\,\Delta T^r$ the expression $\mathrm{e}^{\alpha(t-s)}$ is bounded, and formula (11) follows from integration of $p(s)$ and from the use of $\Delta T \le t$ where necessary.

We remark that a direct proof of the estimate (11) without using the ideas of backward error analysis is given in [6], where also the dependence of the constant c_k on k is considered. The present proof has the advantage that a slight modification yields sharp estimates of the energy error for Hamiltonian systems.

Harmonic oscillator. We consider the harmonic oscillator with Hamiltonian (5), and we apply symplectic integrators for $\varphi^F_{\Delta T}$ and $\varphi^G_{\Delta T}$. In this situation we have

$$\left\|\frac{\partial y}{\partial y_0}\Big(t, s, u^k(s)\Big)\right\| \le \hat{c}, \qquad \|u^0(t) - y_F(t)\| \le c_0\,t\,\Delta T^r, \qquad (27)$$

which are the estimates (25) with $L = 0$. We thus obtain

$$\|u^k(t) - y_F(t)\| \le c_k\,(t\,\Delta T^r)^{(k+1)},$$

which yields the estimates (6) of Theorem 1 without specifying the constant c_k.

Integrable differential equations. We either consider integrable (or nearly integrable) Hamiltonian systems and apply symplectic integrators for $\varphi^F_{\Delta T}$ and $\varphi^G_{\Delta T}$, or we consider integrable reversible systems (see [7, Chapter XI]) and apply reversible (symmetric) numerical methods. Then we have

$$\left\|\frac{\partial y}{\partial y_0}\Big(t, s, u^k(s)\Big)\right\| \le \hat{c}\,(1+(t-s)), \qquad \|u^0(t) - y(t)\| \le c_0\,t\,\Delta T^r. \qquad (28)$$

The induction argument above gives the estimate

$$\left\|g^k_{\Delta T}(\mathbf{u}^k(t))\right\| \le C_k\Big(\|u^k(t) - y_F(t)\| + p(t)\,\Delta T^{rk}\Big) \qquad (29)$$

with $p(t) = c_{k-1}\,(1+\Delta T^{r+1})\,t^{2k-1} + c_{k-2}\,(1+\Delta T^{r+1})\,t^{2k-3}\Delta T + \ldots + c_0\,t\,\Delta T^{k-1}$, and the function $m(t) = \|u^k(t) - y_F(t)\|$ is seen to satisfy

$$m(t) \le \alpha \int_0^t \Big(1 + (t-s)\Big)\Big(m(s) + p(s)\,\Delta T^{rk}\Big)\,\mathrm{d}s, \qquad \alpha = \hat{c}\,C_k \Delta T^r.$$

In this situation, the techniques of proof for Gronwall's inequality[2] lead to

$$m(t) \le \frac{\alpha}{\beta} \int_0^t \mathrm{e}^{\alpha(t-s)/2} \sinh(\beta\,(t-s))\,(\dot{p}(s) + p(s))\,\Delta T^{rk}\,\mathrm{d}s$$

with $\beta = \sqrt{\alpha + \alpha^2/4}$. For bounded $t^2\Delta T^r$ (which means that $t\beta$ and $t\alpha$ are bounded) we have $\mathrm{e}^{\alpha(t-s)/2}\sinh(\beta\,(t-s)) \le C\beta\,(t-s)$ for $0 \le s \le t$. The estimate (8) of Theorem 2 is then obtained by using $t \ge 1$ after integration.

6.2. Estimates of the energy error

After having obtained estimates for the global error of the parareal algorithm, we are interested in the long-time energy preservation which can explain the numerical observations of Section 3. For the harmonic oscillator the long-time behavior is completely understood with the analysis of Section 4.

We consider a Hamiltonian system $\dot{y} = J^{-1}\nabla H(y)$, and we apply symplectic integrators for $\varphi^F_{\Delta T}$ and $\varphi^G_{\Delta T}$. The modified differential equation of the fine integrator is then also Hamiltonian with $H_F(y)$ of (4), and that of the parareal iterates is given by

$$\dot{u}^k = J^{-1}\nabla H_F(u^k) + \Delta T^r g^k_{\Delta T}(\mathbf{u}^k).$$

When computing the total derivative of $H_F(u^k(t))$ the leading term cancels due to the skew-symmetry of J, so that the fundamental theorem of calculus leads to

$$H_F(u^k(t)) - H_F(y_0) \;=\; \Delta T^r \int_0^t \nabla H_F(u^k(s))^{\mathsf{T}}\, g^k_{\Delta T}(\mathbf{u}^k(s))\,\mathrm{d}s. \tag{30}$$

Comparing this relation with (24) we notice that instead of the (linearly or exponentially) growing resolvent we have the bounded gradient of H_F. This is the reason why the exponent of t is typically reduced by one in the error of the energy. Using the estimates for the global error together with Corollary 5 yields information on the near energy preservation.

Integrable differential equations. We consider integrable problems for which (28) holds. For bounded $t^2\Delta T^r$ it follows from the estimate (8) that the term $\|u^k(t) - y_F(t)\|$ in (29) can be neglected and we obtain the estimate

$$\Big|H_F(u^k(t)) - H_F(y_0)\Big| \;\le\; c_k\,t^{2k}\Delta T^{r(k+1)} \;=\; c_k\,\Delta T^r(t^2\Delta T^r)^k.$$

This perfectly agrees with the numerical experiments illustrated in Figure 2.

[2] Define $u(t) = \int_0^t (t-s)(m(s) + p(s)\Delta T^{rk})\,\mathrm{d}s$, so that $m(t) \le \alpha(\dot{u}(t) + u(t))$ and $\ddot{u}(t) = \alpha\dot{u}(t) + \alpha u(t) + f(t)$ with $f(t) \le p(t)\Delta T^{rk}$. The solution of this differential equation is $u(t) = \beta^{-1}\int_0^t \mathrm{e}^{\alpha(t-s)/2}\sinh(\beta(t-s))\,f(s)\,\mathrm{d}s$.

Chaotic Hénon–Heiles problem. For this problem we do not have better estimates than (25) and (26). Again the exponent of t is reduced by one, so that

$$\left|H_F(u^k(t)) - H_F(y_0)\right| \;\le\; c_k\, t^k\, \mathrm{e}^{Lt}\, \Delta T^{r(k+1)}.$$

The presence of the factor e^{Lt} implies that the length of the interval, where the parareal iterations can improve the numerical approximations, is more or less independent of ΔT. This agrees well with the observation in Figure 3.

7. Symmetric parareal algorithm

A standard approach for making a one-step method symmetric is to compose it with its adjoint, i.e., $\Psi_{\Delta T/2} \circ \Psi^*_{\Delta T/2}$, where $\Psi^*_{\Delta T/2} = \Psi^{-1}_{-\Delta T/2}$. For the parareal algorithm (17) this has been proposed in [5]. In terms of the discrete flows $\varphi^F_{\Delta T}$ and $\varphi^G_{\Delta T}$ the *symmetric parareal algorithm* is given by

$$\begin{aligned}
u^0_{n+1/2} &= (\varphi^G_{-\Delta T/2})^{-1}(u^0_n) \\
u^0_{n+1} &= \varphi^G_{\Delta T/2}(u^0_{n+1/2}) \\
u^k_{n+1/2} &= \left(\varphi^G_{-\Delta T/2}\right)^{-1}\left(u^k_n - \varphi^F_{-\Delta T/2}(u^{k-1}_{n+1/2}) + \varphi^G_{-\Delta T/2}(u^{k-1}_{n+1/2})\right) \\
u^k_{n+1} &= \varphi^F_{\Delta T/2}(u^{k-1}_{n+1/2}) - \varphi^G_{\Delta T/2}(u^{k-1}_{n+1/2}) + \varphi^G_{\Delta T/2}(u^k_{n+1/2}).
\end{aligned} \tag{31}$$

The modified differential equation of an adjoint method is obtained by changing the sign of ΔT, and the modified differential equation of the composition of two integrators is obtained by the Baker–Campbell–Hausdorff formula, see [7, Section IX.4]. Assume that both methods, $\varphi^F_{\Delta T}$ and $\varphi^G_{\Delta T}$, are symmetric. The modified differential equation of the symmetric parareal algorithm is then given by

$$\dot u^k = f_{\Delta T}(u^k) + 2^{1-r}\Delta T^r g^{k,S}_r(u^{k-1}, u^k) + \Delta T^{r+2} g^{k,S}_{r+2}(u^{k-1}, u^k) + \ldots, \tag{32}$$

where $g^{k,S}_r$ is the same function as g^k_r in (20). The expansion is in even powers of ΔT, so that there is no ΔT^{r+1} term. However, the dominating perturbation term in the equation (32) is still not Hamiltonian.

Analysis for the harmonic oscillator. We use the notation $G_+ = G(\Delta T/2)$, $G_- = G(-\Delta T/2)$, and similarly for $F(\Delta T)$. With a generating function $u_n(\zeta) = \sum_{k\ge 0} u^k_n\, \zeta^k$, the recursion for the symmetric parareal algorithm can be written as

$$u_{n+1}(\zeta) = \Big(G_+ + \zeta\,(F_+ - G_+)\Big)\Big(G_- + \zeta\,(F_- - G_-)\Big)^{-1} u_n(\zeta) = \Big(G_+G_- + \sum_{l\ge 1} \zeta^l\, D_l\Big)\, u_n(\zeta)$$

Notice that from $\zeta = 1$ we obtain $G_+G_- + \sum_{l\ge 1} D_l = F_+F_-$. Furthermore, we have $D_l = O(\Delta T^{l(r+1)})$ if $F(\Delta T) - G(\Delta T) = O(\Delta T^{r+1})$.

The same analysis as for the classical parareal algorithm yields

$$u^k_n = (F_+F_-)^n y_0 - \sum_{l\ge k+1} \sum_{j=1}^{l} \binom{n}{j} (G_+G_-)^{n-j} \sum_{l_1+\ldots+l_j=l} D_{l_1} \cdot \ldots \cdot D_{l_j}\, y_0.$$

The dominant term in the sum is for $l = k+1$ and $j = l$. Since $D_1 = F_+F_- - G_+G_- = O(\Delta T^{2(r+1)})$, this term is the same as in (14) with $F(\Delta T)$ and $G(\Delta T)$ replaced by F_+F_- and G_+G_-, respectively. Consequently, the same conclusions can be drawn as before, and we obtain that the classical and symmetric parareal algorithms give more or less identical results. This perfectly agrees with the numerical experiments presented in [5].

8. Derivative parareal algorithm for accurate long-time integrations

The parareal algorithm is a parallel technique that permits to achieve high accuracy (that of the fine integrator) with a reduced cost. One learns in the undergraduate studies that for achieving high accuracy in an efficient way, one needs a numerical integrator of high order. One of the main messages of the present article is that also the coarse integrator should have a certain accuracy. Otherwise the convergence of the parareal iterations would be too slow, and the time window, where the algorithm can be applied, would be rather small preventing an efficient integration. At this point we would like to emphasize that the purpose of the computations of Section 3 is the illustration of the theoretical results; in practice, one should never use the second order Störmer–Verlet method, neither as fine integrator nor as coarse integrator.

Choice of the fine integrator. Implicit Runge–Kutta method (collocation based on Gauss-Legendre quadrature) with $s = 6$ stages. It has order $2s = 12$, it is symmetric and symplectic, and thus has all favorable properties for the integration of Hamiltonian systems.

Choice of the coarse integrator. Implicit Runge–Kutta method (collocation based on Gauss-Legendre quadrature) with $s = 3$ or $s = 4$ stages. Also a composition method of order 6, based on the Störmer–Verlet method, could be considered.

With such a choice the use of the parareal algorithm is questionable. Assuming that the cpu time of a Runge–Kutta step is proportional to the number of stages, the fine integrator will be not more than twice as expensive as the coarse integrator. Even if one solves the nonlinear Runge–Kutta equation for the coarse integrator with an accuracy that increases with the number of iterations (taking care that convergence is not affected), the fine integrator will typically be not more than 4 times more expensive.

A notable exception is when very high accuracy is required (for example in the simulation of planetary motion), so that a computation in quadruple precision could be considered for the fine integrator, keeping the computations of the coarse integrator with the much cheaper double precision. However, there is a serious obstacle. Although the correction $\varphi^G_{\Delta T}(u^k_n) - \varphi^G_{\Delta T}(u^{k-1}_n)$ in (3) is very small, the cancellation of significant digits in performing this difference prevents to get more accuracy than with a computation in double precision. We therefore propose a modification of (3) that circumvents this difficulty.

We note that the just mentioned difference can be approximated by

$$\varphi^G_{\Delta T}(u^k_n) - \varphi^G_{\Delta T}(u^{k-1}_n) \approx D\varphi^G_{\Delta T}(u^k_n)(u^k_n - u^{k-1}_n),$$

where D means the derivative with respect to the initial value (this is one step back in the derivation of the parareal algorithm as presented in [6]). This derivative can be efficiently computed by applying the method $\varphi^G_{\Delta T}$ to the augmented system

$$\begin{pmatrix} \dot y \\ \dot z \end{pmatrix} = \begin{pmatrix} f(y) \\ f'(y)\, z \end{pmatrix}, \qquad \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} u^k_n \\ u^k_n - u^{k-1}_n \end{pmatrix}. \tag{33}$$

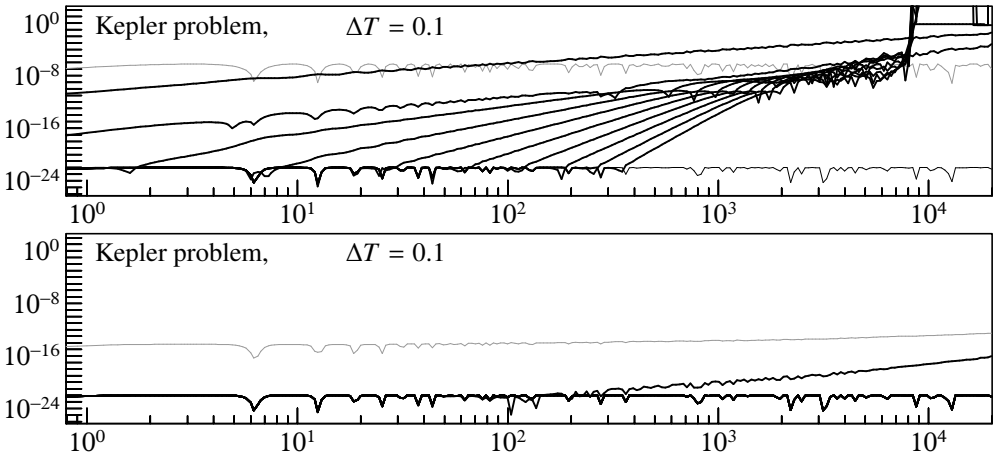

Figure 4: Error in the Hamiltonian for the modified parareal algorithm applied to the Kepler problem.; fine integrator is the quadruple implementation of the Gauss-Runge–Kutta method of order 12, the coarse integrator is 21 steps of the Störmer–Verlet method (upper pricture) and the Gauss–Runge–Kutta method of order 8 (lower pircture), both in double precision implementation.

In fact, the numerical outcome is $y_1 = \varphi^G_{\Delta T}(u^k_n)$ (not needed) and $z_1 = D\varphi^G_{\Delta T}(u^k_n)(u^k_n - u^{k-1}_n)$. Using the ideas of automatic differentiation the evaluation of the augmented vector field (33) is typically less than twice as expensive[3] than that of $f(y)$ alone.

This discussion motivates the consideration of the following modification, which we call *derivative parareal algorithm*:

$$\begin{aligned} u^0_{n+1} &= \varphi^G_{\Delta T}(u^0_n) \\ u^k_{n+1} &= \varphi^F_{\Delta T}(u^{k-1}_n) + D\varphi^G_{\Delta T}(u^k_n)(u^k_n - u^{k-1}_n). \end{aligned} \tag{34}$$

The values $\varphi^F_{\Delta T}(u^{k-1}_n)$ can still be computed in parallel. Let us denote c'_G the cost (cpu time) of one step of $\varphi^G_{\Delta T}$ applied to (33). We typically have $c'_G < 2c_G$. The total time of the standard application of the algorithm is then $Nc_G + K(Nc'_G + c_F)$ with N processors, and for the pipelined version it is $Nc'_G + Kc_F$ with $(q+1)K + 1$ processors. These times have to be compared with those of Section 2.

We apply this algorithm to the Kepler problem using data as in Figure 2. With $\Delta T = 0.1$ the 12th order Gauss–Runge–Kutta method (in quadruple precision) as fine integrator yields an error in the Hamiltonian that is approximately 10^{-22}. The nonlinear Runge–Kutta equations are solved by fixed-point iterations. We exploit the fact that the Hamiltonian is separable, so that a factor ΔT^2 is gained in each iteration. To achieve an accuracy of 10^{-26} it takes 8 fixed-point iterations (that is $8\times6+1 = 49$ force evaluations per step; one of them is used for starting approximations). Note that the vectors u^k_n have to be stored in quadruple precision. In the experiment of Figure 4 we use two different methods for the coarse integrator (both implemented in double precision). In the lower picture of Figure 4 we use the Gauss–Runge–Kutta method of order 8, which takes 5 fixed-point iterations (that is $5 \times 4 + 1 = 21$ force evaluations per step) to achieve an accuracy

[3]For N-body problems the computation of the square roots takes an important amount of time. This computation needs not be repeated for the lower component of (33).

of 10^{-12}. In the upper picture of Figure 4 we use 21 steps of the Störmer–Verlet scheme with step size $\Delta T/21$ for $\varphi^G_{\Delta T}$. The number 21 is chosen, so that the coarse integrator takes the same number of function evaluations in both situations.

We see that the algorithm (34) yields a precision far beyond that of double precision arithmetic, even though the coarse integrator is applied only in double precision. This experiment clearly demonstrates the importance of the use of coarse integrators that are sufficiently accurate. In the upper picture the accuracy of the coarse integrator (method of order 2) is approximately 10^{-7} which is not enough. One needs 4 parareal iterations to cover one period of the solution. Moreover, it is not possible to apply the parareal algorithm to time windows of length ≥ 400. In the lower picture the accuracy of the coarse integrator (method of order 8) is about 10^{-14}. This time one gets the accuracy of the fine integrator after one iteration on the interval of length 300, and after 2 iterations on an interval of length larger than 20 000. Under the assumption that a computation in quadruple precision is 7 times more expensive than the same computation in double precision, we have that $c_F \approx (7 \cdot 49/21)\, c_G$. This is large compared to c'_G, so that an important improvement of a parallel computation can be expected.

References

[1] C. Audouze, M. Massot, S. Volz, Symplectic multi-time step parareal algorithms applied to molecular dynamics, hal.archives-ouvertes.fr/hal-00358459 (2009) 18 pages.

[2] G. Bal, Q. Wu, Symplectic parareal, in: Domain decomposition methods in science and engineering XVII, volume 60 of *Lect. Notes Comput. Sci. Eng.*, Springer, Berlin, 2008, pp. 401–408.

[3] A. Bellen, M. Zennaro, Parallel algorithms for initial value problems for difference and differential equations, J. Comput. Appl. Math. 25 (1989) 341–350.

[4] P. Chartier, B. Philippe, A parallel shooting technique for solving dissipative ODEs, Computing 51 (1993) 209–236.

[5] X. Dai, C. Le Bris, F. Legoll, Y. Maday, Symmetric parareal algorithms for Hamiltonian systems, ESAIM Math. Model. Numer. Anal. (2012).

[6] M.J. Gander, E. Hairer, Nonlinear convergence analysis for the parareal algorithm, in: Domain decomposition methods in science and engineering XVII, volume 60 of *Lecture Notes in Computational Science and Engineering*, Springer, Berlin, 2008, pp. 45–56.

[7] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations, Springer Series in Computational Mathematics 31, Springer-Verlag, Berlin, 2nd edition, 2006.

[8] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I. Nonstiff Problems, Springer Series in Computational Mathematics 8, Springer, Berlin, 2nd edition, 1993.

[9] P.J. van der Houwen, B.P. Sommeijer, CWI contributions to the development of parallel Runge-Kutta methods, Appl. Numer. Math. 22 (1996) 327–344. Special issue celebrating the centenary of Runge-Kutta methods.

[10] H. Jiménez-Pérez, J. Laskar, A time-parallel algorithm for almost integrable Hamiltonian systems, In Proceedings of CoRR (2011) 19 pages.

[11] J.L. Lions, Y. Maday, G. Turinici, Résolution d'EDP par un schéma en temps "pararéel", C. R. Acad. Sci. Paris Sér. I Math. 332 (2001) 661–668.

[12] M.L. Minion, A hybrid parareal spectral deferred corrections method, Commun. Appl. Math. Comput. Sci. 5 (2010) 265–301.