On the Superlinear and Linear Convergence of the Parareal Algorithm

Martin J. Gander[1] and Stefan Vandewalle[2]

[1] Section de Mathématiques, University of Geneva, Switzerland
Martin.Gander@math.unige.ch
[2] Dept. of Computer Science, Leuven, Belgium
Stefan.Vandewalle@cs.kuleuven.be

Summary. The parareal algorithm is a method to solve time dependent problems parallel in time: it approximates parts of the solution later in time simultaneously to parts of the solution earlier in time. In this paper the relation of the parareal algorithm to space-time multigrid and multiple shooting methods is first briefly discussed. The focus of the paper is on some new convergence results that show superlinear convergence of the algorithm when used on bounded time intervals, and linear convergence for unbounded intervals.

1 Introduction

The parareal algorithm was first presented in [LMT01] to solve evolution problems in parallel. The name was chosen to indicate that the algorithm is well suited for parallel real time computations of evolution problems whose solution can not be obtained in real time using one processor only. The method approximates successfully the solution later in time before having fully accurate approximations from earlier times. The algorithm has received a lot of attention over the past few years; for extensive experiments and studies of convergence and stability issues we refer to [MT02, FC03] and the contributions in the 15th Domain Decomposition Conference Proceedings [KHP$^+$04].

Parareal is not the first algorithm to propose the solution of evolution problems in a time-parallel fashion. Already in 1964, Nievergelt suggested a parallel time integration algorithm [Nie64], which led to multiple shooting methods. The idea is to decompose the time integration interval into subintervals, to solve an initial value problem on each subinterval concurrently, and to force continuity of the solution branches on successive intervals by means of a Newton procedure. Since then, many variants of the method were developed and used for the time-parallel integration of evolution problems, see e.g. [BZ89, CP93]. In [GV05], we show that the parareal algorithm can be interpreted as a particular multiple shooting method, where the Jacobian matrix is approximated in a finite difference way on the coarse mesh in time.

In 1967, Miranker and Liniger [ML67] proposed a family of predictor-corrector methods, in which the prediction and correction steps can be performed in parallel over a number of time-steps. Their idea was to "widen the computational front", i.e., to allow processors to compute solution values on several time-steps concurrently. A similar motivation led to the block time integration methods by Shampine and Watts [SW69]. More recently, [SN88] and [Wom90] considered the time-parallel application of iterative methods to the system of equations derived with implicit time-integration schemes. Instead of iterating until convergence over each time step before moving on to the next, they showed that it is possible to iterate over a number of time steps at once. Thus a different processor can be assigned to each time step and they all iterate simultaneously. The acceleration of such methods by means of a multigrid technique led to the class of parabolic multigrid methods, as introduced in [Hac84]. The multigrid waveform relaxation and space-time multigrid methods also belong to that class. In [VdV94], a time-parallel variant was shown to achieve excellent speedups on a computer with 512 processors; while run as sequential algorithm the method is comparable to the best classical time marching schemes. Experiments with time-parallel methods on 2^{14} processors are reported in [HVW95]. In [GV05], it is shown that the parareal algorithm can also be cast into the parabolic multigrid framework. In particular, parareal can be identified with a two level multigrid Full Approximation Scheme, with a special Jacobi-type smoother, with strong semi-coarsening in time, and selection and extension operators for restriction and interpolation.

2 A Review of the Parareal Algorithm

The parareal algorithm for the system of ordinary differential equations

$$\boldsymbol{u}' = \boldsymbol{f}(\boldsymbol{u}), \quad \boldsymbol{u}(0) = \boldsymbol{u}_0, \quad t \in [0,T], \tag{1}$$

is defined using two propagation operators. Operator $G(t_2, t_1, \boldsymbol{u}_1)$ provides a rough approximation to $\boldsymbol{u}(t_2)$ of the solution of (1) with initial condition $\boldsymbol{u}(t_1) = \boldsymbol{u}_1$, whereas operator $F(t_2, t_1, \boldsymbol{u}_1)$ provides a more accurate approximation of $\boldsymbol{u}(t_2)$. The algorithm starts with an initial approximation $\boldsymbol{U}_k^0$, $k = 0, 1, \ldots, \bar{k}$ at the time points $t_0, t_1, \ldots, t_{\bar{k}}$ given for example by the sequential computation of $\boldsymbol{U}_{k+1}^0 = G(t_{k+1}, t_k, \boldsymbol{U}_k)$, with $\boldsymbol{U}_0 = \boldsymbol{u}_0$, and then performs for $n = 0, 1, 2, \ldots$ the correction iteration

$$\boldsymbol{U}_{k+1}^{n+1} = G(t_{k+1}, t_k, \boldsymbol{U}_k^{n+1}) + F(t_{k+1}, t_k, \boldsymbol{U}_k^n) - G(t_{k+1}, t_k, \boldsymbol{U}_k^n). \tag{2}$$

Note that, for $n \to \infty$, the method will upon convergence generate a series of values $\boldsymbol{U}_k$ that satisfy $\boldsymbol{U}_{k+1} = F(t_{k+1}, t_k, \boldsymbol{U}_k)$. That is, the approximation at the time-points t_k will have achieved the accuracy of the F-propagator. Alternatively, one can restrict the number of iterations of (2) to a finite value. In that case, (2) defines a new time-integration scheme. The accuracy of the

$\boldsymbol{U}_k^n$ values is characterized by a theorem from [LMT01]. The theorem applies for a scalar linear problem of the form

$$u' = -au, \quad u(0) = u_0, \quad t \in [0, T]. \tag{3}$$

Theorem 1. *Let $\Delta T = T/\bar{k}$, $t_k = k\Delta T$ for $k = 0, 1, \ldots, \bar{k}$. Let $F(t_{k+1}, t_k, U_k^n)$ be the exact solution at t_{k+1} of (3) with $u(t_k) = U_k^n$, and $G(t_{k+1}, t_k, U_k^n)$ the corresponding backward Euler approximation with time step ΔT. Then,*

$$\max_{1 \le k \le \bar{k}} |u(t_k) - U_k^n| \le C_n \Delta T^{n+1}, \tag{4}$$

where the constant C_n is independent of ΔT.

Hence, for a fixed iteration step n, the algorithm behaves in ΔT like a method $O(\Delta T^{n+1})$. Note that the convergence of the algorithm for a fixed ΔT and increasing number of iterations n is not covered by the above theorem, because the constant C_n is growing in n in the estimate of the proof in [LMT01].

3 Convergence analysis for a scalar ODE

We show two new convergence result for fixed ΔT when n becomes large. The first result is valid on bounded time intervals, $T < \infty$, whereas the second one also holds for unbounded time intervals. The results apply for an arbitrary explicit or implicit one step method applied to (3) with $a \in \mathbb{C}$, i.e., $U_{k+1} = \beta U_k$, in the region of absolute stability of the method, i.e., $|\beta| \le 1$.

In our analysis an important role will be played by a strictly upper triangular Toeplitz matrix M of size $\bar{k}$. Its elements are defined as follows,

$$M_{ij} = \begin{cases} \beta^{j-i-1} & \text{if } j > i, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

A key property of M, whose proof we omit here, is that

$$|\beta| \le 1 \quad \Longrightarrow \quad ||M^n||_\infty \le \binom{\bar{k}-1}{n}. \tag{6}$$

Theorem 2 (Superlinear convergence on bounded intervals). *Let $T < \infty$, $\Delta T = T/\bar{k}$, and $t_k = k\Delta T$ for $k = 0, 1, \ldots, \bar{k}$. Let $F(t_{k+1}, t_k, U_k^n)$ be the exact solution at t_{k+1} of (3) with $u(t_k) = U_k^n$, and let $G(t_{k+1}, t_k, U_k^n) = \beta U_k^n$ be a one step method in its region of absolute stability, i.e., $|\beta| \le 1$. Then,*

$$\max_{1 \le k \le \bar{k}} |u(t_k) - U_k^n| \le \frac{|e^{-a\Delta T} - \beta|^n}{n!} \prod_{j=1}^{n} (\bar{k} - j) \max_{1 \le k \le \bar{k}} |u(t_k) - U_k^0|. \tag{7}$$

If the local truncation error of G is bounded by $C\Delta T^{p+1}$, then

$$\max_{1 \le k \le \bar{k}} |u(t_k) - U_k^n| \le \frac{(CT)^n}{n!} \Delta T^{pn} \max_{1 \le k \le \bar{k}} |u(t_k) - U_k^0|. \tag{8}$$

Proof. We denote by e_k^n the error at iteration step n of the parareal algorithm at time t_k, $e_k^n := u(t_k) - U_k^n$. With (2) and an induction argument on k, it is easy to see that this error satisfies

$$e_k^{n+1} = \beta e_{k-1}^{n+1} + (e^{-a\Delta T} - \beta) e_{k-1}^n = (e^{-a\Delta T} - \beta) \sum_{j=1}^{k-1} \beta^{k-j-1} e_j^n.$$

This relation can be written in matrix form by collecting e_k^n in the vector $\boldsymbol{e}^n = (e_{\bar{k}}^n, e_{\bar{k}-1}^n, \ldots, e_1^n)^T$, which leads to

$$\boldsymbol{e}^{n+1} = (e^{-a\Delta T} - \beta) M \boldsymbol{e}^n, \tag{9}$$

where matrix M is given in (5). By induction on (9), we obtain

$$||\boldsymbol{e}^n||_\infty \le |(e^{-a\Delta T} - \beta)|^n ||M^n||_\infty ||\boldsymbol{e}^0||_\infty, \tag{10}$$

which together with (6) implies (7). The bound (8) follows from the bound on the local truncation error together with a simple estimate of the product,

$$\frac{|e^{-a\Delta T} - \beta|^n}{n!} \prod_{j=1}^{n} (\bar{k} - j) \le \frac{C^n \Delta T^{(p+1)n}}{n!} \bar{k}^n = \frac{(CT)^n}{n!} \Delta T^{pn}.$$

Remark 1. The product term in (7) shows that the parareal algorithm converges for any ΔT on any bounded time interval in at most $\bar{k} - 1$ steps. Furthermore the algorithm converges superlinearly, as the division by $n!$ in (7) shows. Finally, if instead of an exact solution on the subintervals a fine grid approximation is used, the proof remains valid with some minor modifications.

Theorem 3 (Linear convergence on long time intervals). *Let ΔT be given, and $t_k = k\Delta T$ for $k = 0, 1, \ldots$. Let $F(t_{k+1}, t_k, U_k^n)$ be the exact solution at t_{k+1} of (3) with $u(t_k) = U_k^n$, and let $G(t_{k+1}, t_k, U_k^n) = \beta U_k^n$ be a one step method in its region of absolute stability, with $|\beta| < 1$. Then,*

$$\sup_{k>0} |u(t_k) - U_k^n| \le \left(\frac{|e^{-a\Delta T} - \beta|}{1 - |\beta|} \right)^n \sup_{k>0} |u(t_k) - U_k^0|. \tag{11}$$

If the local truncation error of G is bounded by $C\Delta T^{p+1}$, then

$$\sup_{k>0} |u(t_k) - U_k^n| \le \left(\frac{C\Delta T^p}{\Re(a) + O(\Delta T)} \right)^n \sup_{k>0} |u(t_k) - U_k^0|. \tag{12}$$

Proof. In the present case M, as defined in (5), is an infinite dimensional Toeplitz operator. Its infinity norm is given by

$$||M||_\infty = \sum_{j=0}^{\infty} |\beta|^j = \frac{1}{1 - |\beta|}.$$

Using (9), we obtain for the error vectors $\boldsymbol{e}^n$ of infinite length the relation

$$||\boldsymbol{e}^n||_\infty \leq |(e^{-a\Delta T} - \beta)|^n ||M||_\infty^n ||\boldsymbol{e}^0||_\infty = \left(\frac{|(e^{-a\Delta T} - \beta)|}{1-|\beta|}\right)^n ||\boldsymbol{e}^0||_\infty, \quad (13)$$

which proves the first result. For the second result, the bound on the local truncation error, $|e^{-a\Delta T} - \beta| \leq C\Delta T^{p+1}$, implies for $p > 0$ that $\beta = 1 - a\Delta T + O(\Delta T^2)$, and hence $1 - |\beta| = \Re(a)\Delta T + O(\Delta T^2)$, which implies (12).

4 Convergence analysis for partial differential equations

We now use the results derived in Section 3 to investigate the performance of the parareal algorithm on partial differential equations. We consider two model problems, a diffusion problem and an advection problem. For the diffusion case, we consider the heat equation, without loss of generality in one dimension,

$$u_t = u_{xx}, \quad \text{in } \Omega = \mathbb{R}, \quad u(0,x) \in L^2(\Omega). \quad (14)$$

Using a Fourier transform in space, this equation becomes a system of decoupled ordinary differential equations for each Fourier mode ω,

$$\hat{u}_t = -\omega^2 \hat{u}, \quad (15)$$

and hence the convergence results of Theorems 2 and 3 can be directly applied. If we discretize the heat equation in time using backward Euler, then we have the following convergence result for the parareal algorithm.

Theorem 4 (Heat Equation Convervence Result). *Under the conditions of Theorem 2, if* $a = \omega^2$*, and* $G(t_{k+1}, t_k, U_k^n) = \beta U_k^n$ *with* $\beta = \frac{1}{1+\omega^2\Delta T}$ *from backward Euler, the parareal algorithm has a superlinear bound on the convergence rate on bounded time intervals,*

$$\max_{1\leq k\leq \bar{k}} ||u(t_k) - U_k^n||_2 \leq \frac{\gamma_s^n}{n!} \prod_{j=1}^{n} (\bar{k} - j) \max_{1\leq k\leq \bar{k}} ||u(t_k) - U_k^0||_2, \quad (16)$$

where $||\cdot||_2$ *denotes the spectral norm in space and the constant* γ_s *is universal,* $\gamma_s = 0.2036321888$*. On unbounded time intervals, we have*

$$\sup_{k>0} ||u(t_k) - U_k^n||_2 \leq \gamma_l^n \sup_{k>0} ||u(t_k) - U_k^0||_2, \quad (17)$$

where the universal constant $\gamma_l = 0.2984256075$.

Proof. A simple calculation shows that the numerator in the superlinear bound (7) is for backward Euler uniformly bounded,

$$|e^{-\omega^2\Delta T} - \frac{1}{1+\omega^2\Delta T}| \leq \gamma_s,$$

where the maximum γ_s is attained at $\omega^2 \Delta T = \bar{x}_s := 2.512862417$. This leads to (16) by using the Parseval-Plancherel identity.

The convergence factor in the linear bound (12) is also bounded,

$$\frac{|e^{-\omega^2 \Delta T} - \frac{1}{1+\omega^2 \Delta T}|}{1 - \frac{1}{1+\omega^2 \Delta T}} \leq \gamma_l,$$

where the maximum γ_l is attained at $\omega^2 \Delta T = \bar{x}_l := 1.793282133$, which leads to (17) using the Parseval-Plancherel identity.

Next, we consider a pure advection problem

$$u_t = u_x, \quad \text{in } \Omega = \mathbb{R}, \quad u(0,x) \in L^2(\Omega). \tag{18}$$

Using a Fourier transform in time, this equation becomes

$$\hat{u}_t = -i\omega \hat{u}. \tag{19}$$

The convergence results of Theorems 2 and 3 can be directly applied. If we discretize the advection equation in time using backward Euler, then we have the following convergence result for the parareal algorithm.

Theorem 5 (Advection Equation Convergence Result). *Under the conditions of Theorem 2, if $a = -i\omega$, and $G(t_{k+1}, t_k, U_k^n) = \beta U_k^n$ with $\beta = \frac{1}{1+i\omega \Delta T}$ from backward Euler, the parareal algorithm has a superlinear bound on the convergence rate on bounded time intervals,*

$$\max_{1 \leq k \leq \bar{k}} ||u(t_k) - U_k^n||_2 \leq \frac{\alpha_s^n}{n!} \prod_{j=1}^{n} (\bar{k} - j) \max_{1 \leq k \leq \bar{k}} ||u(t_k) - U_k^0||_2, \tag{20}$$

where the constant α_s is universal, $\alpha_s = 1.224353426$.

Proof. A simple calculation shows that the numerator in the superlinear bound (7) is for backward Euler uniformly bounded,

$$|e^{-i\omega \Delta T} - \frac{1}{1 + i\omega \Delta T}| \leq \alpha_s,$$

which leads to (20) using the Parseval-Plancherel identity.

Remark 2. There is no long term convergence result for (18). The convergence factor in (11) is not bounded by a quantity less than one.

5 Numerical Experiments

In order to verify the theoretical results, we first show some numerical experiments for the scalar model problem (3) with $f = 0$, $a = 1$, $u_0 = 1$. The

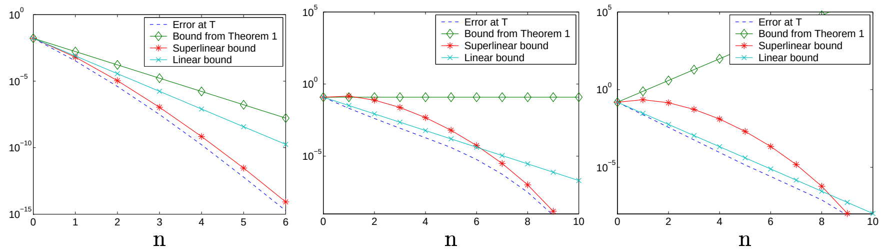

Fig. 1. Parareal convergence for (3) on a short, medium and long time interval.

Backward Euler method is chosen for both the coarse approximation and the fine approximation, with time step ΔT and $\Delta T/\bar{m}$ respectively. We show in Figure 1 the convergence results obtained for $T = 1$, $T = 10$ and $T = 50$, using $\bar{k} = 10$ and $\bar{m} = 20$ in each case. One can clearly see that parareal has two different convergence regimes: for $T = 1$, the algorithm converges superlinearly, and the superlinear bound from Theorem 2 is quite sharp. For $T = 10$, the convergence rate is initially linear, and then a transition occurs to the superlinear convergence regime. Finally, for $T = 50$, the algorithm is in the linear convergence regime and the bound from Theorem 3 is quite sharp. Note also that the bound from Theorem 1 indicates stagnation for $T = 10$, since $\Delta T = 1$, and divergence for $T = 50$, since then $\Delta T > 1$. The parareal algorithm does however also converge for $\Delta T \geq 1$.

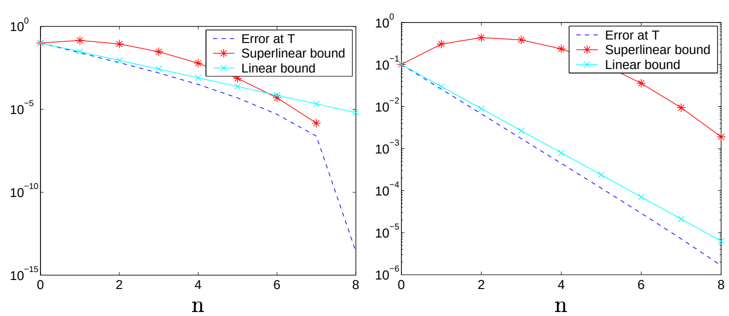

Fig. 2. Error in the L^∞ norm in time and L^2 norm in space for the parareal algorithm applied to the heat equation, on a short (left) and long (right) interval.

We now turn our attention to the PDE case and show some experiments for the heat equation $u_t = u_{xx} + f$, in $(0, L) \times (0, T]$ with homogeneous initial and boundary conditions and with $f = x^4(1 - x) + t^2$. The domain length L is chosen such that the linear bound in (17) of Theorem 4 is attained, which implies that $L = \pi\sqrt{\Delta T/\bar{x}_s}$. With $\Delta T = 1/2$ and $\bar{m} = 10$ we obtain the results shown in Figure 2. On the left, results are shown for $T = 4$, where the algorithm with $\Delta T = 1/2$ will converge in 8 steps. One can see that this is clearly the case. Before that, the algorithm is in the superlinear convergence regime, as predicted by the superlinear bound. Note that the latter bound indicates zero as the error at the eighth step, and thus can not be plotted on the logarithmic scale. On the right, the error is shown for $T = 8$, and the algorithm is clearly in the linear convergence regime.