Richardson extrapolation for space-time discretization methods with application to the advection equation

István Faragó¹², Ágnes Havasi¹³*, and Zahari Zlatev⁴

¹MTA-ELTE Numerical Analysis and Large Networks Research Group
Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary

²Institute of Mathematics, Budapest University of Technology and Economics
Egry J. u. 1., H-1111 Budapest, Hungary

³Department of Applied Analysis and Computational Mathematics
Eötvös Loránd University
Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary

⁴Department of Environmental Science – Atmospheric Modeling
Aarhus University
Frederiksborgvej 399, 4000 Roskilde, Denmark

*Corresponding author E-mail: havasia@cs.elte.hu

(Manuscript received in final form July 8, 2018)

Abstract—Richardson extrapolation is a numerical procedure which enables us to enhance the accuracy of any convergent numerical method in a simple and powerful way.

In this paper we overview the theoretical background of Richardson extrapolation in space and time, where two numerical solutions, obtained on a coarse and a fine space-time grid are combined by a suitable weighted average. We show that when the Crank-Nicolson method is appropriately combined with this extrapolation technique for the solution of the one-dimensional advection equation, then the order of accuracy increases by two both in time and space. The theoretically derived consistency order and the necessity of the smoothness conditions for the exact solution and for the advection velocity are illustrated by numerical experiments, performed by the advection module of the Danish Eulerian Model (DEM).

Key-words: Richardson extrapolation, convergence order, truncation error, advection equation, Crank–Nicolson method
1. Introduction

L.F. Richardson proposed extrapolation as a powerful tool to accelerate the convergence of a sequence of approximations (see Richardson, 1911, 1927). The main idea is to apply the same convergent numerical algorithm, depending on some parameter, by using two different values of this parameter, and combine the obtained numerical solutions to eliminate the leading error term and achieve better accuracy. Its main area of application is the numerical solution of time-dependent ordinary differential equations (see e.g., Marchuk and Shaidurov 1983; Bulirsch and Stoer, 1966; Bader and Deuflhard, 1983, Zlatev et al., 2017). In this case the problem is solved by the same numerical method using two different time-step sizes, and the obtained numerical solutions are combined at each time layer of the coarser grid. In this manner, under certain smoothness conditions, the order of the numerical scheme increases by one. (We remark that by using more than two numerical solutions, even higher order accuracy can be achieved.) In Faragó et al. (2013) it is proved that the combination of any diagonally implicit Runge–Kutta method with active Richardson extrapolation is consistent and zero-stable, and consequently it is convergent. In Zlatev et al. (2012), the stability of the $\theta$-method combined with the active Richardson extrapolation is investigated on a fixed mesh, and among others it is shown that the combined method is strongly A-stable for the values of $\theta$ in the range $[2/3, 1]$. The extrapolated solution is only calculated at the nodes of the coarse grid. Roache and Knupp (1993) extended the technique so that the extrapolated solution can also be calculated at the fine grid points. Interesting details about some of the scientists, who initiated the work on different extrapolation methods (including here the Richardson extrapolation) can be found in Brezinski and Redivo Zaglia (1993).

The idea of Richardson extrapolation can be applied during the numerical solution of partial differential equations (PDEs) as well. In the majority of the applications, the equations are first discretized in space, which results in a system of time-dependent ordinary differential equations (ODEs), and the Richardson extrapolation is applied with the aim of increasing the order of the time integration method only (see e.g., Havasi et al., 2013; Mona et al., 2015). However, it is also possible to apply this technique with spatial and temporal step sizes refined simultaneously, and the numerical solutions obtained on the coarse and fine space-time grids are then combined with appropriate weight coefficients. This procedure for the case of one spatial dimension has been elaborated in Richards (1997). In this paper we apply the idea of Richardson extrapolation in space and time to the one-dimensional advection equation, when the underlying space-time discretization method is the Crank–Nicolson scheme.

The structure of the paper is as follows. In Section 2, we present the general idea of Richardson extrapolation as a convergence acceleration method
and go into more details about the application of this idea during the numerical solution of ODEs. In Section 3, the Richardson extrapolation in space and time is presented for the numerical solution of PDEs with one spatial dimension as proposed by Richards (1977), who specified the condition of a proper grid refinement. In Section 4, this idea is applied to the one-dimensional advection equation when solved by the Crank–Nicolson scheme as an underlying method. We will see that in this particular application, the gain in the accuracy is even better than expected, if certain smoothness conditions are satisfied by the exact solution of the advection problem and by the advection velocity. Our theoretical results are illustrated with three numerical experiments, performed by the advection module of UNI-DEM (the Unified Danish Eulerian Model).

2. The principle of Richardson extrapolation

Consider a numerical algorithm depending on parameter \( \tau \), which tends to the exact value \( A^* \) when \( \tau \) tends to zero. Typically, \( \tau \) denotes the mesh size of the discretization. If the method is convergent to the order \( p \), then by solving the problem with the parameter values \( \tau_1 = \tau \) and \( \tau_2 = \tau/2 \), the equalities

\[
A^* - A(\tau) = K \tau^p + O(\tau^{p+1}),
\]

and

\[
A^* - A(\tau/2) = K(\tau/2)^p + O(\tau^{p+1}),
\]

hold, where the quantity \( K \) is independent of the value of the parameter \( \tau \). During Richardson extrapolation, the two numerical solutions are combined by suitable weight coefficients according to the formula

\[
A_{RE}(\tau) = -\frac{1}{2^p - 1} A(\tau) + \frac{2^p}{2^p - 1} A(\tau/2) = \frac{2^p A(\tau/2) - A(\tau)}{2^p - 1},
\]

see Richardson (1911, 1927).

This procedure is rather general, and it can be applied in combination with any convergent numerical method. It is primarily used during the numerical solution of Cauchy problems for ordinary differential equations, where \( A(\tau) \) denotes the numerical solution of the problem by using time step \( \tau \). Consider the Cauchy problem

\[
\begin{align*}
y'(t) &= f(t, y), & t &\in [0, T] \\
y(0) &= y_0,
\end{align*}
\]
where \( f : \mathbb{R}^{M+1} \to \mathbb{R}^M \), \( y_0 \in \mathbb{R}^M \) is a given initial vector, and the unknown function \( y \) is of type \( \mathbb{R} \to \mathbb{R}^M \). Let us define the following two grids on the time interval \([0, T]\):

\[
\Omega_\tau := \{ t_n = n \tau, n = 0, 1, \ldots, N \},
\]

and

\[
\Omega_{\tau/2} := \{ t_k = k \tau / 2, k = 0, 1, \ldots, 2N \},
\]

with \( N ; \tau = T \). Denote by \( z(t_n) \) the numerical solution obtained at time \( t_n \) of the coarse grid, and \( w(t_n) \) that obtained at the same time layer \( t_n \) on the fine grid. (Note that in the second case twice as many steps are taken than in the first case.) Then, following the principle of Richardson extrapolation, the approximation

\[
y_{RE}(t_n) := \frac{2^p w(t_n) - z(t_n)}{2^p - 1}
\]

provides a numerical solution, accurate to the order \( p+1 \), in case the exact solution \( y(t) \) is \( p+1 \) times continuously differentiable. During the so-called passive Richardson extrapolation, the combined solution is never used during the further computations, while in the case of the active Richardson extrapolation always the combined solution is propagated, both on the coarse grid and on the fine grid.

**3. Richardson extrapolation in space and time**

The above idea can be extended to PDEs depending on time and on one space dimension as follows. Two space-time grids are defined on the solution domain \( 0 \leq x \leq L, 0 \leq t \leq T \) of the problem. Denote by \( \Delta x \) and \( \Delta t \) the spatial and temporal step sizes of the fine grid, respectively. On the coarse grid, the step sizes \( m\Delta x \) and \( m^\beta \Delta t \) will be used by certain integers \( m \geq 1 \) and \( \beta \geq 1 \), such that both grids cover the solution domain.

One can prove by using some simplifying assumptions (Richards, 1997) that the exponent \( \beta \) needs to satisfy a certain requirement in order that higher order of accuracy can be achieved, provided that the exact solution is sufficiently smooth. Let us fix a grid point \( \tilde{x} \) which is an element of both the fine and the coarse spatial grid, and denote the truncation error of the method at point \( (\tilde{x}, t_n) = (\tilde{x}, n\Delta t) \) by \( E_T(\tilde{x}, t_n, \Delta x, \Delta t) \). For a numerical method that has order \( p_x \) in space and order \( p_t \) in time, the truncation error can be given as.
\[ E_T(\vec{x}, t_n, \Delta x, \Delta t) = O((\Delta x)^{p_x}, (\Delta t)^{p_t}) = K_x(\Delta x)^{p_x} + K_t(\Delta t)^{p_t} + O((\Delta x)^{q_x}, (\Delta t)^{q_t}). \]  

Here \( q_x \) and \( q_t \) are integers for which \( q_x > p_x \) and \( q_t > p_t \), and \( K_x \) and \( K_t \) depend on the problem and on the solution, but are independent of \( \Delta x \) and \( \Delta t \). One can show that on a time interval of length \( \Delta t \), an error of the size \( \Delta t E_T \) arises. Denote by \( u(\vec{x}, t) \) the exact solution of the problem at the point \((\vec{x}, t)\), by \( y^n_f \) the numerical solution on the fine grid, and by \( y^n_c \) the numerical solution on the coarse grid, both at the point \((\vec{x}, n\Delta t)\). Then the error arising during the computation of \( u \) on the fine grid, by the assumption that we use the exact solution at time \( n\cdot\Delta t \), can be written as

\[ u(\vec{x}, (n + 1)\Delta t) - y^{n+1}_f = K_x \Delta t(\Delta x)^{p_x} + K_t(\Delta t)^{p_t+1} + O(\Delta t(\Delta x)^{q_x}, (\Delta t)^{q_t+1}). \]  

The error of the coarse grid solution after a time step of length \( m^\beta \Delta t \) is

\[
\begin{align*}
&u(\vec{x}, (n + m^\beta)\Delta t) - y^{n+m^\beta}_c \\
&= m^\beta \Delta t(\Delta x)^{p_x} K_x + (m^\beta \Delta t)^{p_t+1} K_t + O(\Delta t(\Delta x)^{q_x}, (\Delta t)^{q_t+1}) \\
&= m^{p_x+\beta} \Delta t(\Delta x)^{p_x} K_x + m^{\beta(p_t+1)}(\Delta t)^{p_t+1} K_t + O(\Delta t(\Delta x)^{q_x}, (\Delta t)^{q_t+1}).
\end{align*}
\]  

The error on the fine grid at time \((n + m^\beta)\Delta t\) can be expanded as follows, by assuming that the errors of successive time steps add up, moreover, that the coefficients \( K_x \) and \( K_t \) do not change:

\[
\begin{align*}
&u(\vec{x}, (n + m^\beta)\Delta t) - y^{n+m^\beta}_f \\
&= m^\beta \Delta t(\Delta x)^{p_x} K_x + m^\beta(\Delta t)^{p_t+1} K_t + O((\Delta t)^2(\Delta x)^{q_x}, \Delta t(\Delta x)^{q_x}, (\Delta t)^{p_t+2}, (\Delta t)^{q_t+1}).
\end{align*}
\]  

During Richardson extrapolation in space and time, a linear combination of the numerical solutions \( y^{n+m^\beta}_c \) and \( y^{n+m^\beta}_f \) is determined at the points of the coarse grid as

\[ y^{n+m^\beta}_{RE} = \gamma_1 y^{n+m^\beta}_c + \gamma_2 y^{n+m^\beta}_f, \]  

where \( \gamma_1 \) and \( \gamma_2 \) are real constants, to be chosen appropriately. By substitution from Eqs. (10) and (11) we are led to the equality

\[
\begin{align*}
y^{n+m^\beta}_{RE} &= (\gamma_1 + \gamma_2)u(\vec{x}, (n + m^\beta)\Delta t) \\
&+ (\gamma_1 m^\beta + \gamma_2 m^{p_t+1})\Delta t(\Delta x)^{p_x} K_x + (\gamma_1 m^\beta + \gamma_2 m^{\beta(p_t+1)})(\Delta t)^{p_t+1} K_t \\
&+ O((\Delta t)^2(\Delta x)^{p_x}, \Delta t(\Delta x)^{q_x}, (\Delta t)^{p_t+2}, (\Delta t)^{q_t+1}).
\end{align*}
\]
Our aim is to eliminate the leading error terms, containing the coefficients $K_x$ and $K_t$. Obviously, this can only be achieved if

$$\gamma_1 + \gamma_2 = 1; \quad \gamma_1 m^\beta + \gamma_2 m^{p_x + \beta} = 0, \quad (14)$$

and

$$\gamma_1 m^\beta + \gamma_2 m^{(p_x + 1)} = 0. \quad (15)$$

Subtracting (16) from (15), one can easily see that

$$\gamma_2 (m^{p_x + \beta} - m^{(p_x + 1)}) = 0 \quad (17)$$

should hold, which in case of $\gamma_2 \neq 0$ can only be satisfied if $p_x + \beta = \beta (p_t + 1)$, i.e., if $\beta = p_x / p_t$.

Thus, it has been shown that provided the underlying space-time discretization method has order $p_x$ in space and order $p_t$ in time, then, for the increase of the order of accuracy, $\beta$ should be chosen such that it is equal to the ratio of $p_x$ and $p_t$. The convergence order of the method obtained in this manner is increased to $O((\Delta t)^2 (\Delta x)^{p_x}, \Delta t(\Delta x)^{q_x}, (\Delta t)^{p_t + 1}, (\Delta t)^{q_t + 1})$, which corresponds to a method with truncation error of $O((\Delta t)^2 (\Delta x)^{p_x}, (\Delta x)^{q_x}, (\Delta t)^{p_t + 1}, (\Delta t)^{q_t + 1})$.

4. Application to the 1D advection equation

In this section we present an application of the above approach. Consider the one-dimensional advection equation

$$\partial_t u = -c \partial_x u, \quad x \in [0, L], \quad t \in [t_0, T] \quad (18)$$

with appropriate initial and boundary conditions, where $c = c(x, t)$. Assume that we want to solve this problem with the Crank–Nicolson method

$$\frac{u_{i}^{n+1} - u_{i}^{n}}{\Delta t} = \frac{1}{2} \left( -c_{i}^{n+0.5} \frac{u_{i+1}^{n+1} - u_{i-1}^{n+1}}{2\Delta x} \right) + \frac{1}{2} \left( -c_{i}^{n+0.5} \frac{u_{i+1}^{n} - u_{i-1}^{n}}{2\Delta x} \right). \quad (19)$$

As it is known, the Crank–Nicolson method has second order both in space and time, i.e., $p_x = p_t = 2$ (Strikwerda, 2004). According to the previous section, the proper choice of $\beta$ is $\beta = p_x / p_t = 1$. Consequently, if the stepsize of the coarse grid is twice the stepsize $\Delta x$ of the fine grid (i.e., $m = 2$), then $\Delta t$ should also by
multiplied by 2 to generate the coarse time grid. Further, we need the coefficients $\gamma_1$ and $\gamma_2$. From condition (15), by substitution of the values $m = 2$, $p_x = 2$ and $\beta = 1$, the equation

$$\gamma_1 + 4\gamma_2 = 0$$

is obtained, from which, in view of Eq. (14), we get the coefficients $\gamma_1 = -1/3$, and $\gamma_2 = 4/3$. (Note that these weights are the same as those applied when Richardson extrapolation is performed only in time for a method of order 2.) Moreover, since $\Delta t = \Delta x$, therefore by Eq. (13) one can expect that the order of accuracy increases by one, both in space and time. However, in this case we get an even better acceleration, as it is shown by direct expansion into Taylor series in Zlatev et al., 2011: If the advection velocity $c = c(x, t)$ is twice continuously differentiable both by $x$ and $t$, and the exact solution $u(x, t)$ is four times continuously differentiable both by $x$ and $t$, then the combination of the Crank–Nicolson method with the active Richardson extrapolation has order four both in space and time for problem Eq. (18). (For details of the proof see the given reference.)

4.1. Numerical experiments

Our theoretical results were tested by using different parameter choices and initial conditions together with periodic boundary conditions in the advection problem Eq. (18). Note that in order that the active Richardson extrapolation can be applied, at all time layers of the coarse grid interpolation is needed to the fine grid so that we have the combined solution on the fine grid before the next time step.

4.2. Test problem 1

In the first experiment performed by the advection module of UNI-DEM, the following parameter values were specified: $L = 50000000$, $t_0 = 43200$, $T = 129600$, $c(x, t) = 320$, and we used the initial function

$$u_0(x) = 1.4679 \cdot 10^{12} (1 + 99 e^{-10^{-12}(x-1000000)^2})$$

This is a smooth function with sharp gradients, see Fig. 1.
The exact solution has the form \( u(x, t) = u_0(x - 320(t - 43200)) \), which has the same smoothness property as the initial function. Table 1 shows the global errors obtained by spatial and time steps, halved several times simultaneously. When the Crank–Nicolson method is applied alone, the error is reduced by a factor of four to a good approximation, i.e., the method behaves as a fourth-order one, as expected. Moreover, when combined with the active Richardson extrapolation, the errors roughly decrease by a factor of 16, so the theoretically obtained fourth-order convergence is obtained.

**Table 1.** Global errors obtained in Test problem 1 by the Crank–Nicolson method alone (CN) and by its combination with the active Richardson extrapolation (CN + RE). \( N_t \): the number of time steps; \( N_x \): the number of spatial steps. In parentheses the factors are given by which the global error decreased in comparison with the error in the previous row.

<table>
<thead>
<tr>
<th>( N_t )</th>
<th>( N_x )</th>
<th>CN</th>
<th>CN + RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>168</td>
<td>160</td>
<td>7.37E-01</td>
<td>1.454E-01</td>
</tr>
<tr>
<td>336</td>
<td>320</td>
<td>4.003E-01 (1.842)</td>
<td>1.741E-02 (8.350)</td>
</tr>
<tr>
<td>672</td>
<td>640</td>
<td>1.254E-01 (3.142)</td>
<td>1.224E-03 (14.220)</td>
</tr>
<tr>
<td>1344</td>
<td>1280</td>
<td>3.080E-02 (4.135)</td>
<td>7.730E-05 (15.837)</td>
</tr>
<tr>
<td>2688</td>
<td>2560</td>
<td>7.765E-03 (3.967)</td>
<td>4.841E-06 (15.970)</td>
</tr>
<tr>
<td>5376</td>
<td>5120</td>
<td>1.954E-03 (3.974)</td>
<td>3.026E-07 (15.999)</td>
</tr>
<tr>
<td>10752</td>
<td>10240</td>
<td>4.892E-04 (3.994)</td>
<td>1.891E-08 (16.004)</td>
</tr>
<tr>
<td>21504</td>
<td>20480</td>
<td>1.224E-04 (3.999)</td>
<td>1.181E-09 (16.011)</td>
</tr>
</tbody>
</table>
4.3. Test problem 2

In this experiment, the parameters were chosen as $L = 2\pi$, $t_0 = 0$, $T = 2\pi$, $c(x, t) = 0.5$, and the initial function was the highly oscillatory function shown in Fig. 2:

$$u_0(x) = 1.4679 \cdot 10^{12} (100 + 99 \sin(10x)).$$

Fig. 2. The initial function in Test problem 2.

The exact solution is $u(x, t) = u_0(x - 0.5t)$.

The global errors are shown in Table 2. With the Crank–Nicolson method applied alone, similar results were obtained as before. However, when combined with the active Richardson extrapolation, the convergence is slower than fourth (but faster than third) order. The reason for this is probably the fact that the third-order interpolation polynomial, applied for interpolation of the coarse grid solution to the fine grid, cannot be applied at the boundary of the space domain, so here a second-order formula was used.
Table 2. Global errors obtained in Test problem 2 by the Crank–Nicolson method alone (CN) and by its combination with the active Richardson extrapolation (CN + RE). $N_t$: the number of time steps; $N_x$: the number of spatial steps. In parentheses the factors are given by which the global error decreased in comparison with the error in the previous row.

<table>
<thead>
<tr>
<th>$N_t$</th>
<th>$N_x$</th>
<th>CN</th>
<th>CN + RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>168</td>
<td>160</td>
<td>7.851E-01</td>
<td>1.560E-02</td>
</tr>
<tr>
<td>336</td>
<td>320</td>
<td>2.160E-01 (3.635)</td>
<td>1.227E-03 (12.713)</td>
</tr>
<tr>
<td>672</td>
<td>640</td>
<td>5.317E-02 (4.062)</td>
<td>1.072E-04 (11.432)</td>
</tr>
<tr>
<td>1344</td>
<td>1280</td>
<td>1.327E-02 (4.007)</td>
<td>1.150E-05 (9.333)</td>
</tr>
<tr>
<td>2688</td>
<td>2560</td>
<td>3.319E-03 (3.997)</td>
<td>1.193E-06 (9.641)</td>
</tr>
<tr>
<td>5376</td>
<td>5120</td>
<td>8.299E-04 (4.000)</td>
<td>1.478E-07 (8.071)</td>
</tr>
<tr>
<td>10752</td>
<td>10240</td>
<td>2.075E-04 (4.000)</td>
<td>1.618E-08 (9.136)</td>
</tr>
<tr>
<td>21504</td>
<td>20480</td>
<td>5.187E-05 (4.000)</td>
<td>1.965E-09 (8.233)</td>
</tr>
</tbody>
</table>

4.4. Test problem 3

In this case, the parameters of the first test problems were used with the exception that the initial function was changed to the following one:

$$u_0(x) = \begin{cases} 
1.4679 \cdot 10^{12} & \text{if } x \leq 5000000 \text{ or } x \geq 15000000 \\
1.4679 \cdot 10^{12} \left(1 + 99 \frac{x - 5000000}{5000000}\right) & \text{if } 5000000 < x \leq 10000000 \\
1.4679 \cdot 10^{12} \left(1 + 99 \frac{15000000 - x}{5000000}\right) & \text{if } 10000000 < x \leq 15000000 
\end{cases}$$

see Fig. 3. This initial function is continuous, but its first derivative is discontinuous, and so none of the sufficient conditions for the fourth-order convergence of the combined method is satisfied. As Table 3 shows, the theoretically derived order is not obtained for any of the two methods, however, the Richardson extrapolated method is still more accurate.
Fig. 3. The initial function in Test problem 3.

<table>
<thead>
<tr>
<th>$N_t$</th>
<th>$N_x$</th>
<th>CN</th>
<th>CN + RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>168</td>
<td>160</td>
<td>1.353E-01</td>
<td>4.978E-02</td>
</tr>
<tr>
<td>336</td>
<td>320</td>
<td>7.687E-02 (1.760)</td>
<td>2.761E-02 (1.803)</td>
</tr>
<tr>
<td>672</td>
<td>640</td>
<td>4.424E-02 (1.737)</td>
<td>1.551E-02 (1.780)</td>
</tr>
<tr>
<td>1344</td>
<td>1280</td>
<td>2.555E-02 (1.732)</td>
<td>8.570E-03 (1.810)</td>
</tr>
<tr>
<td>2688</td>
<td>2560</td>
<td>1.636E-02 (1.561)</td>
<td>4.590E-03 (1.867)</td>
</tr>
<tr>
<td>5376</td>
<td>5120</td>
<td>1.051E-02 (1.552)</td>
<td>2.318E-03 (1.980)</td>
</tr>
<tr>
<td>10752</td>
<td>10240</td>
<td>5.551E-03 (1.899)</td>
<td>1.188E-03 (1.951)</td>
</tr>
<tr>
<td>21504</td>
<td>20480</td>
<td>2.921E-03 (1.900)</td>
<td>6.575E-04 (1.807)</td>
</tr>
</tbody>
</table>

Table 3. Global errors obtained in Test problem 3 by the Crank–Nicolson method alone (CN) and by its combination with the active Richardson extrapolation (CN + RE). $N_t$: the number of time steps; $N_x$: the number of spatial steps. In parentheses the factors are given by which the global error decreased in comparison with the error in the previous row.
5. Conclusion

Richardson extrapolation is a powerful tool to accelerate the convergence of numerical methods, and it can be applied not only to ordinary differential equations, but also to partial differential equations. In the latter case we either use the extrapolation after space discretization, or we apply Richardson extrapolation in space and time. In this paper we presented how the latter method works on the one-dimensional advection equation when the underlying numerical scheme is the Crank–Nicolson method. Our numerical experiments confirmed the theoretically derived order increase, and showed that satisfying the smoothness conditions is important to get this high order convergence.

References