Stability of implicit extrapolation methods

Ulrich Rüde[1]

1 Introduction

Multilevel methods decompose the solution space in a nested sequence of subspaces. These subspaces are then used to construct a multigrid or multilevel preconditioned conjugate gradient method. Thus the multilevel structure is the basis for the efficient solution of a partial differential equation (PDE). In this setting, the solution is defined in the topmost space and the multilevel structure is just used to accelerate some iterative solution method.

Besides this *algebraic* perspective, the multilevel structure may be also used to improve the accuracy of the discretization itself. Under certain conditions, the nested mesh and space structure can be exploited by various *extrapolation* schemes.

Extrapolation results for finite difference methods can be found in [MS83] and for finite elements in [BLR86]. All *explicit* extrapolation techniques rely on the existence of *global error expansions* for the approximate solutions, typically of the form

$$u_h - u = \sum_{i=1}^{k} h^{\alpha_i} e_i + R_{k+1}, \tag{1}$$

where u_h and u are the numerical and the true solution of the differential equation, e_i are functions independent of h, and R_{k+1} is a remainder term. The parameter h denotes the mesh width and can be interpreted as identifying a single space V_h in the nested multilevel structure. Once an expansion of the form (1) has been proven to

[1] Institut für Mathematik, Universität Augsburg, D-86135 Augsburg, Germany, email: ruede@math.uni-augsburg.de

Domain Decomposition Methods in Sciences and Engineering, edited by R. Glowinski *et al.*

exist, and when the exponent α_1 of the leading term has been identified, then a linear combination of u_h and u_{2h} of the form

$$\bar{u}_h = \frac{2^{\alpha_1}}{2^{\alpha_1}-1}u_h - \frac{1}{2^{\alpha_1}-1}u_{2h}$$

will lead to an approximation where the dominating error term is eliminated. Using this scheme recursively, leads to the well-known *Richardson extrapolation table.*

Extrapolation is only computationally feasible, when the coefficients α_i are known, and when they form a quickly growing sequence. If the α_1 do not grow quickly, so that the expansion has many terms of almost equal order, extrapolation will be less attractive, since many linear combinations are needed before the approximation order is significantly increased.

Unfortunately, these situations are typical in many practical PDE problems when reentrant corners, interfaces with jumps in the coefficients or rough data may lead to *singularities*, each of which may create its own error contributions to the error expansion with a new set of exponents α_i.

In the case of reentrant corner singularities, the α_i are still known and a local extrapolation technique has been proposed in [BR88], based on meshes refined locally according to the interior angles.

In this paper we will discuss another variant of extrapolation, where the Richardson principle is applied indirectly. Such *implicit extrapolation* methods have been introduced in [JR94, Rüd91a]. They are closely related to the so-called τ-extrapolation in multigrid methods, see e.g. [Hac85]Chapter 14.1.3.

Implicit extrapolation is based on a local element-by-element analysis of numerical quadrature and differentiation rules. Thus they primarily depend on the smoothness of the shape functions rather than the global regularity of the solution. The regularity of the solution is of course required to justify the use of high order (polynomial) approximations in general. Implicit extrapolation has the advantage that it is formally not applied to the solution itself, but only to the finite element functions approximating the solution. Thus the specific analysis is independent of the solution properties. Whether high order finite elements are suitable to approximate a given solution can be decided independently. Furthermore, implicit extrapolation is suitable for nonuniform grids as they may occur in a local refinement context.

Using extrapolation in a multilevel context is especially attractive, when the extrapolation to obtain higher order discretizations is efficiently integrated with a fast multilevel solver. For integrating Richardson extrapolation with a full multigrid method and a comparison with multigrid τ-extrapolation, see e.g. Lin and Schüller [SL85]. The implicit extrapolation which is the main topic of this paper is most efficiently implemented by the multigrid τ-extrapolation algorithm. Due to the space limitations in this paper we refer for all algorithmic details to [JR94, Rüd91a]. We remark only that implicit extrapolation can be implemented by a trivial change of a conventional (FAS) multigrid method which consists of only a multiplication by an additional extrapolation factor in the fine-to-coarse transfer. Thus integrating implicit extrapolation into an existing multilevel algorithm is easy, and the basic solver remains untouched, since the higher order accuracy is obtained by a defect correction-like iteration.

2 The implicit extrapolation method

For exposition, we consider the simplest case of an elliptic PDE in one dimension

$$u'' = f, \quad u(0) = u(1) = 0. \tag{2}$$

We will use the equivalent formulation as a minimization problem

$$\min_{u \in H_0^1(0,1)} E(u), \quad \text{where } E(u) = \int_0^1 (u'(x))^2 - 2u(x)f(x)\,dx, \tag{3}$$

where $H_0^1(0,1)$ denotes the the usual Sobolev space of order 1 enforcing homogeneous Dirichlet boundary conditions. Next, we introduce the mesh $0 = x_0 < x_1 < \ldots < x_n = 1$, and discretize (3) directly, by representing the continuous function u by the vector $u_h = (u_0^h, \ldots, u_n^h)^T$. With the mesh widths $h_i = x_{i+1} - x_i$ and the midpoints $x_{i+1/2} = 1/2(x_{i+1} + x_i)$, we may replace

$$u'(x_{i+1/2}) \approx \frac{u(x_{i+1}) - u(x_i)}{h_i}$$

and the integration by the *midpoint sum*

$$\int_0^1 f(x)\,dx \approx \sum_{i=0}^{n-1} h_i f(x_{i+1/2})$$

for the numerical approximation to $(u'(x))^2$ and the *trapezoidal sum*

$$\int_0^1 f(x)\,dx \approx \sum_{i=0}^{n-1} \frac{h_i}{2}(f(x_{i+1}) + f(x_i))$$

for approximating the integral over $u(x)f(x)$. These approximations combined, lead to

$$E(u) \approx E_h(u_h) = \sum_{i=0}^{n-1} h_i \left[\frac{1}{2} \left(\frac{u_{i+1} - u_i}{h_i} \right)^2 + \frac{f_{i+1}u_{i+1} + f_i u_i}{2} \right].$$

The corresponding normal equations are

$$\begin{aligned} u_0 &= 0 \\ \frac{-u_{i+1} + u_i}{h_{i+1}} + \frac{u_i - u_{i-1}}{h_i} + \frac{h_{i+1} + h_i}{2} f_i &= 0 \quad \text{for } i = 1, \ldots, N-1 \\ u_N &= 0 \end{aligned}$$

Thus we have recovered the discretization by second order finite differences, or, equivalently, the discretization by piecewise linear finite elements with lumped mass matrix.

The basic reason for deriving the discrete system in this form is the existence of asymptotic error expansions for

$$E(u) - E_h(u) = e_2 h^2 + \ldots + e_{2k} h^{2k} + R_{2k+1} \tag{4}$$

when u is sufficiently smooth. This result holds even for a nonuniform basic discretization $x_0, \ldots, x_n$, if only the refined grids are constructed by recursively inserting the interval midpoints. This result is proved in [Lyn68]. Based on the expansion (4) we may now consider extrapolated functionals of the form

$$\bar{E}_h(u_h) = \frac{4}{3}E_h(u_h) - \frac{1}{3}E_{2h}(u_h), \quad \bar{\bar{E}}_h(u_h) = \frac{16}{15}\bar{E}_h(u_h) - \frac{1}{15}\bar{E}_{2h}(u_h), \quad \text{etc.} \tag{5}$$

Here we interpret u_h simply as a vector of values, and note that $E_{2h}(u_h)$ only uses every second value in u_h.

Obviously, each of the extrapolated functionals defines a new system of normal equations for u_h. In section 3 we will give conditions, when the higher order representation of the functional results in an improved accuracy for u_h. Before we continue this argument, we mention that results analogous to (4) hold in two space dimensions, as proved in [Rüd93] for the case of triangular meshes. Of course this is a crucial result, since the main interest here is in methods which generalize to higher space dimensions. In [JR94] it is furthermore shown that the implicit extrapolation method is equivalent to using higher order finite elements in a special case.

3 Stability of implicit extrapolation

In the above section we have introduced the *implicit extrapolation principle* for deriving higher order difference discretizations for elliptic PDE. In this section, we will interpret u_h as a finite element approximation with lowest oder, that is piecewise linear finite elements. By the extrapolation, as in (5), we construct higher order representations of the integrals. These extrapolated quadrature rules are defined for the continued refinement of a basic mesh $x_0, x_1, \ldots, x_n$ by recursively inserting the interval midpoints. If the basic mesh is associated with the space of piecewise linear finite element functions V_h, it would seem natural to identify each of the refined meshes with the corresponding (hierarchical) finite element spaces $V_h \subseteq V_{h/2} \subseteq V_{h/4} \subseteq \ldots$ of piecewise linear functions.

However, since the basic integration/differentiation rules are already correct in all these spaces, their approximation properties cannot be improved by extrapolation. If there is a positive effect of the implicit extrapolation, it cannot be directly understood within the h-refinement space structure.

Therefore, we consider a p-refinement where the basic space of piecewise linear approximations V_h is enlarged by piecewise polynomials of increasing order. Thus the first refinement from space V_h to $\bar{V}_{h/2}$ with a first step of extrapolation corresponds to adding quadratic functions in each interval (x_i, x_{i+1}). The next step to $\bar{V}_{h/4}$ introduces two additional degrees of freedom in each element and this corresponds to adding cubic and 4th order basis functions. In the next level $\bar{V}_{h/8}$ another four degrees of freedom are added and thus all polynomials up to degree 8.

At the same time as we add p-refinement, the extrapolation according to 5 increases the accuracy of the integration. Unfortunately, this increase of accuracy does not keep up with the rapid growth of the spaces. Only in the first step, when quadratic shape functions are added, the extrapolation provides sufficiently accurate integration rules. In the second step of extrapolation, only the functions up to degree 3 are integrated exactly However in this step cubic *and* 4th order functions are added to the solution space. Thus there is a mismatch between the functions in the finite element space and the accuracy of the extrapolated quadrature rules.

In conventional finite element analysis, the need for numerical quadrature to compute the stiffness matrix (and right hand side) is rarely made explicit. The finite element space is basic, and the use of another space (say of nodal values) for numerical quadrature is usually not explicit in the basic analysis. Here, in the context of implicit extrapolation, the converse is true. We directly work in the space used for quadrature and our interpretation of this as a p-version finite element space is just an artifact.

Generally speaking, we are faced with the following situation.

- The continued refinement has produced a large space V which we identify with a p-refinement of our basic discretization. The associated finite element problem is denoted by

$$\min_{u \in V} E(u), \quad \text{where } E(u) = a(u,u) - 2(f,u). \tag{6}$$

- In the space V, however, we cannot assume that we represent the functionals correctly. Our extrapolation procedure falls behind in accuracy, so what we really solve is not (6) but

$$\min_{\tilde{u} \in V} \tilde{E}(\tilde{u}), \quad \text{where } \tilde{E}(\tilde{u}) = \tilde{a}(\tilde{u},\tilde{u}) - 2(f,\tilde{u}) \tag{7}$$

- Though we cannot say, how accurate $\tilde{E}$ is with respect to E in the full space V, we have constructed the method such, that V has a subspace W, where both functionals agree well with each other, say

$$|E(w) - \tilde{E}(w)| \leq \epsilon ||w||^2 \text{ for all } w \in W. \tag{8}$$

In W we consider the equation defined by the original functional

$$\min_{w \in W} E(w) \tag{9}$$

The best we can hope for in this situation is that our computed solution $\tilde{u}$ of problem (7) is close to the solution w of problem (9). This is shown in the following theorem.

Theorem 3.1 (Stability of implicit extrapolation) *Let V be a Hilbert space with norm $||\cdot||$ and bilinear forms $a(\cdot,\cdot)$ and $\tilde{a}(\cdot,\cdot)$, $f \in V^*$ with positive constants c_1, c_2 such that*

$$c_1 ||v||^2 \leq a(u,u) \leq c_2 ||v||^2.$$

Let u, $\tilde{u}$, and w be defined by (6), (7), and (9), respectively. Assume that (8) holds and that

$$\tilde{a}(v,v) \geq a(v,v) \text{ for all } v \in V. \tag{10}$$

Then there exists $c > 0$ such that

$$||w - \tilde{u}|| \leq c(\sqrt{\epsilon}||w|| + ||w - u||).$$

Proof.

$$
\begin{aligned}
\tilde{a}(w-\tilde{u}, w-\tilde{u}) &= \tilde{a}(w, w-\tilde{u}) - \tilde{a}(\tilde{u}, w-\tilde{u}) \\
&= \tilde{a}(w, w-\tilde{u}) - a(u, w-\tilde{u}) \\
&= \tilde{a}(w, w-\tilde{u}) - a(w, w-\tilde{u}) + a(w-u, w-\tilde{u}) \\
&= (\tilde{a}-a)(w, w-\tilde{u}) + a(w-u, w-\tilde{u}) \\
&\leq \sqrt{(\tilde{a}-a)(w,w)\,(\tilde{a}-a)(w-\tilde{u}, w-\tilde{u})} + \sqrt{a(w-u, w-u)\,a(w-\tilde{u}, w-\tilde{u})} \\
&\leq \sqrt{(\tilde{a}-a)(w,w)\,\tilde{a}(w-\tilde{u}, w-\tilde{u})} + \sqrt{a(w-u, w-u)\,\tilde{a}(w-\tilde{u}, w-\tilde{u})}.
\end{aligned}
$$

Therefore

$$\sqrt{\tilde{a}(w-\tilde{u}, w-\tilde{u})} \leq \sqrt{(\tilde{a}-a)(w,w)} + \sqrt{a(w-u, w-u)}.$$

and thus there exists $c > 0$ with

$$||w-\tilde{u}|| \leq c(\sqrt{\epsilon}||w|| + ||w-u||).$$

□

This stability theorem requires condition (10), which states that we must not underestimate the true energy $E(u)$ in our numerical analogue $\tilde{E}(u)$. In many practical situations, this condition is violated as soon as we attempt a second step of extrapolation Thus the potential increase in accuracy is not reflected in the solution.

A computational remedy has been suggested in [Rüd91b]. Obviously it is sufficient so solve (7) subject to the constraint $\tilde{u} \in W$. If such a constraint is used to make V coincide with W, the instability cannot occur.

However, condition (10) indicates that we need not strictly impose $\tilde{u} \in W$, but it suffices to correct a possible underestimate of the energy of solution components which do not lie in W. Since the orthogonal complement of W in V consists of all polynomials of degree larger than some k_0 and smaller than k_1, it is simple to construct such a correction. Thus (7) becomes modified to

$$\min_{\tilde{u} \in V} \tilde{E}(\tilde{u}) + \rho D(\tilde{u}), \tag{11}$$

where ρ is a sufficiently large real parameter and $D(\tilde{u})$ is a quadratic form which vanishes on all (piecewise) polynomials of degree $\leq k_0$ and is positive for all (piecewise) polynomials of degree $> k_0$ and $\leq k_1$. Thus the *accuracy condition* (8) is maintained, and clearly, for some ρ large enough, the *stability condition* (10) will be satisfied, too. Consequently, an ansatz of the form (11) where $\tilde{E}$ is constructed by implicit extrapolation may obtain any approximation order, provided that the *penalty term* D is chosen appropriately.

$D(\tilde{u})$ can e.g. be constructed by computing the finite differences estimating the kth derivatives for $k_0 < k \leq k_1$ in each element. Thus $D(\tilde{u})$ is a local operator, effecting only nodal values within a single element. Computational experiments showing the effectiveness of this technique are given in [Rüd91b].

4 A numerical experiment for a singular interface problem

The implicit extrapolation technique can be generalized to two and three space dimensions. We will now present a two-dimensional numerical experiment in a nontrivial situation. We consider the diffusion equation

$$\begin{aligned} \nabla \cdot \alpha(x,y)\nabla u(x,y) &= 0 \text{ in } (0,1)^2 && (12)\\ u(x,y) &= g(x,y) \text{ on } \partial(0,1)^2, && (13)\end{aligned}$$

where the coefficient has jumps across the lines $x = 1/2$ and $y = 1/2$. In particular, we assume

$$\alpha(x,y) = \begin{cases} 1 & \text{in } (0,1/2)\times(1/2,1)\cup(1/2,1)\times(0,1/2) \\ 3 & \text{in } (0,1/2)^2\cup(1/2,1)^2 \end{cases}.$$

We choose the boundary data such that the solution has the form

$$u(x,y) = u(r,\phi) = \Phi(\phi)r^{(2/3)},$$

where (r,ϕ) are polar coordinates with respect to the point $(x,y) = (1/2,1/2)$, and where $\Phi(\phi)$ is of the form

$$\Phi(\phi) = \sin(2/3\phi + \beta)$$

in each of the four quadrants of the domain, with β chosen such that the interface conditions for the normal derivatives are satisfied. The solution is depicted in Figure 1. Since the singularity is of the same type as in the case of an L-shaped domain,

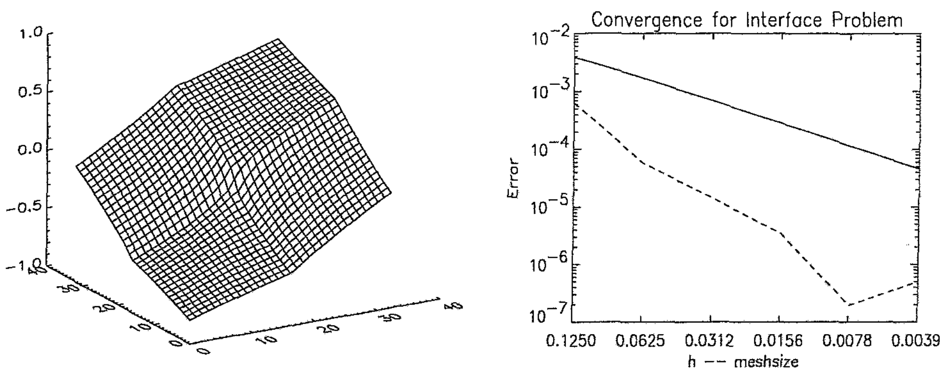

Figure 1 Solution of problem (12,13) and convergence rates.

we expect poor convergence. It has been shown, that in a point-wise sense and for a uniform discretization with piecewise linear triangular elements, we can only expect a convergence rate of $O(h^{4/3})$ This is reflected in the upper (unbroken) line of Figure 1.

A conventional technique to improve on this is to use local refinement to better resolve the singularity. An alternative is based on extrapolation, as in [BR88] or [Rüd88]. This is the approach we will take, focussing on implicit extrapolation.

Thus we solve a system of the form (4), however, we must adapt the extrapolation parameter to the nonstandard leading term in the expansion. We thus minimize $(2^{4/3} - 1)^{-1}(2^{(4/3)}E_h(u_h) - E_{2h}(u_h))$. The convergence rate for the corresponding solution is shown in the lower (broken) line of Figure 1.

The error is visualized in Figure 2. The effect of implicit extrapolation is clearly

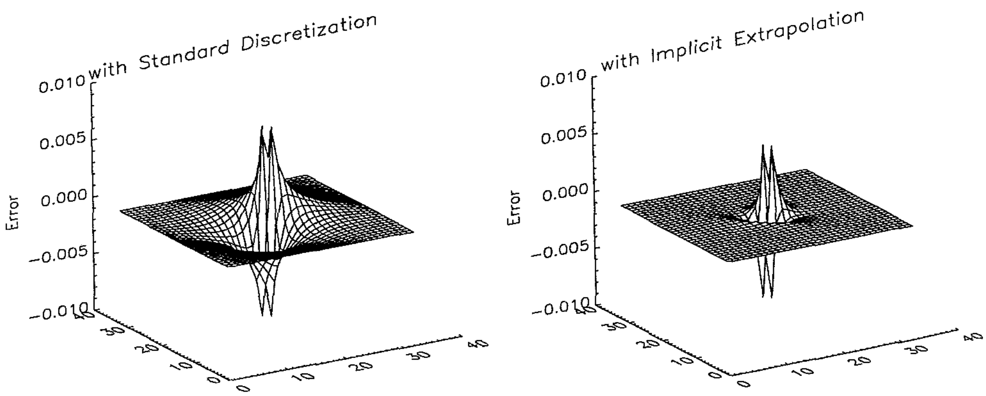

Figure 2 Error for standard discretization and implicit extrapolation.

visible. Since we have not refined locally, the poor resolution of the singularity is not significantly improved. However, with implicit extrapolation, the global spreading of the error, the so-called *pollution effect* is fully suppressed. This is also clear form the convergence graph in Figure 1, where it is shown that implicit extrapolation has recovered $O(h^2)$ convergence away from the singularity.

5 Conclusions

In this paper we have briefly presented the implicit extrapolation method. We have proved a stability condition and have suggested computational techniques to satisfy the stability condition in a general setting. A numerical example has shown, how the method can be used even in the case of singular interface problems. Many more aspects of the method, in particular its combination with splitting extrapolation (see [TmSbL90]) and the so-called sparse grid techniques remain to be analyzed in the future. Some preliminary results are given in [Rüd91b].

References

[BLR86] Blum H., Lin Q., and Rannacher R. (1986) Asymptotic error expansions and Richardson extrapolation for linear finite elements. *Numer. Math.* 49: 11–37.

[BR88] Blum H. and Rannacher R. (1988) Extrapolation techniques for reducing the pollution effect of reentrant corners in the finite element method. *Numer. Math.* 52: 539–564.

[Hac85] Hackbusch W. (1985) *Multigrid Methods and Applications.* Springer Verlag, Berlin.

[JR94] Jung M. and Rüde U. (1994) Implicit extrapolation methods for multilevel finite element computations. In Manteuffel T. (ed) *Preliminary Proceedings of the Colorado Conference on Iterative Methods, Breckenridge, Colorado, April 4-10, 1994.* To appear in S.J.Sci.Comput., 17(1), 1996.

[Lyn68] Lyness J. N. (1968) The calculation of the Stieltjes' integral. *Num. Math.* 12: 252–265.

[MS83] Marchuk G. and Shaidurov V. (1983) *Difference Methods and their Extrapolations.* Springer, New York.

[Rüd88] Rüde U. (1988) On the accurate computation of singular solutions of Laplace's and Poisson's equation. In McCormick S. F. (ed) *Multigrid Methods: Theory, Applications, Supercomputing: Proceedings of the Third Copper Mountain Conference on Multigrid Methods, April 5-10, 1987.* Marcel Dekker, New York, NY, USA.

[Rüd91a] Rüde U. (1991) Adaptive higher order multigrid methods. In Hackbusch W. and Trottenberg U. (eds) *Proceedings of the Third European Conference on Multigrid Methods, October 1-4, 1990,* pages 339–351. Birkhäuser, Boston, MA, USA. International Series of Numerical Mathematics, Vol. 98.

[Rüd91b] Rüde U. (September 1991) Extrapolation and related techniques for solving elliptic equations. Bericht I-9135, Institut für Informatik, TU München.

[Rüd93] Rüde U. (1993) Extrapolation techniques for constructing higher order finite element methods. Bericht I-9304, Institut für Informatik, TU München.

[SL85] Schüller A. and Lin Q. (December 1985) Efficient high order algorithms for elliptic boundary value problems combining full multigrid techniques and extrapolation methods. Arbeitspapiere der GMD 192, Gesellschaft für Mathematik und Datenverarbeitung.

[TmSbL90] Tao L., min Shih T., and bo Liem C. (August 1990) An analysis of splitting extrapolation for multidimensional problems. *Systems Science and Mathematical Sciences* 3(3): 261–272.