Andreas Hohmann

Multilevel Newton h-p Collocation

SC-94-25 (September 1994)

MULTILEVEL NEWTON H-P COLLOCATION

ANDREAS HOHMANN*

Abstract. We derive a simple accuracy matching strategy for inexact Gauss Newton methods and apply it to the numerical solution of boundary value problems of ordinary differential equations by collocation. The matching strategy is based on an affine contravariant convergence theorem, i.e., the characteristic constants are invariant under affine transformations of the domain. The inexact Gauss Newton method is applied to an integral formulation of the BVP. As discretization for the arising linear subproblems we employ adaptive collocation at Gaussian nodes with varying local orders and stepsizes. The grids are chosen via adaptive refinement and order selection.

Key Words. inexact Gauss Newton method, collocation, h-p method, error model,

Introduction. The standard approach to the solution of nonlinear boundary value problems is to first discretize them, e.g. using collocation on a suitable grid, and then to solve the arising finite dimensional nonlinear equation by Newton's method up to the discretization error. This procedure may be iterated using more and more suitable discretizations until the required tolerance is achieved. Due to the mesh independence principle, only very few Newton iterations are necessary on the fine grids, which makes these methods very efficient. This is clearly demonstrated e.g. by the code COLSYS/COLNEW by Ascher, Christiansen, Russell [ACR79] and Bader [Bad88] as the most prominent example.

On the other hand, one could also exchange the order of the two approximation steps and first linearize the problem by Newton's method and then discretize the arising linear subproblems. This approach is often called *quasilinearization* or *multilevel Newton method*. In the context of time dependent problems, these two viewpoints correspond to the method of lines and Rothe's method, respectively. The latter has attracted increasing attention in recent years, which led to highly efficient codes for time dependent problems. We only want to mention the integration methods for parabolic [Bor92] equations and reaction diffusion equations [LW92].

As in the time dependent case, the quasilinearization approach to nonlinear problems appears to be closer to the infinite dimensional nature of the problem, which is reflected by the occurrence of so-called spurious solutions in the standard approach: there may be solutions of the finite dimensional discretized equation which do not correspond to a solution of the original infinite dimensional nonlinear problem. This phenomenon has been observed in particular in the context of continuation methods (see e.g. [Rhe86]).

Applying Newton's method first, it obviously has to be an inexact one, since we cannot solve the infinite dimensional linear subproblems exactly. Hence, we have to think about how to control the accuracy of the linear solvers. To meet the efficiency requirements, we should solve the linear subproblems only as accurately as neccessary for an optimal convergence of Newton's method. Thus, the key is an efficient accuracy matching strategy for the inexact Newton method.

So, this is our program: we derive a simple accuracy matching strategy for inexact Gauss Newton methods and apply it to the numerical solution of boundary value problems of ordinary differential equations by collocation. The matching strategy is based on an affine contravariant convergence theorem, i.e., the characteristic constants are invariant under affine transformations of the domain. The inexact Gauss Newton method is applied to an integral formulation of the BVP. As discretization for the arising linear subproblems we employ adaptive collocation at Gaussian nodes with varying local orders and stepsizes. The grids are chosen via adaptive refinement and order selection.

* Konrad–Zuse–Zentrum Berlin, Heilbronner Str. 10, D-10711 Berlin-Wilmersdorf, Germany. hohmann@sc.zib-berlin.de

1. Inexact Gauss Newton Methods. In this section we tackle the question of how to control the accuracy of the linear subproblems in an inexact Gauss Newton method. We put our emphasis on an implementable accuracy matching strategy that preserves the quadratic convergence of the overall method.

Inexact Newton methods have been attacked by many authors, we only mention Bank and Rose [BR81], Dembo, Eisenstat and Steihaug [DES82], Ypma [Ypm84], Deuflhard [Deu91] and the references therein. Dembo et. al. obtained precise results on how to control the relative residual of the correction equation to obtain a prescribed order $1 < q \leq 2$ of convergence. Unfortunately, they use problem dependent constants, such as bounds for the Jacobian and its inverse, that in many applications cannot be estimated algorithmically or tend to grow with the dimension of the problem (e.g., for successively finer discretizations of nonlinear PDE's).

Deuflhard and Heindl start with a convergence theorem for Newton's method that is invariant with respect to linear transformations of the image space of the nonlinear mapping and therefore only uses the norm in the domain. Due to this *affine invariant* approach, it is possible to give bounds for the relative error of the Newton correction which can be estimated algorithmically. In [Deu91] these methods are extended to the damped (or global) Newton method. The main difficulty of this approach is the control of the relative error which is in many cases much more complicated than the control of the relative residual.

We shall try to combine the advantages of both approaches to obtain algorithmically available bounds for the relative residuals. To this end, we only have to transfer the affine invariant theorems into results which are invariant with respect to affine transformations of the domain and therefore only use the norm given in the image space.

Throughout this chapter we consider a nonlinear problem

$$F(x) = 0$$

where $F : D \subset X \to Y$ is a differentiable mapping from some open subset $D \subset X$ of a Banach space X in another Banach space Y. Newton's method is given by

$$x_{k+1} = x_k + \Delta x_k, \quad F'(x_k)\Delta x_k = -F(x_k) \quad \text{for } k = 0, 1, \ldots,$$

which already includes underdetermined problems solved by a Gauss Newton method. More explicitly, we think of parameter dependent nonlinear problems, where X is the product of the state space and some finite dimensional parameter space.

Substituting approximations s_k for the Newton corrections Δx_k, we obtain an *inexact Newton method*

$$x_{k+1} = x_k + s_k \ , \quad r_k = F'(x_k)s_k + F(x_k)$$

where the *inner residuals* r_k have to be small enough. The accuracy of the approximations s_k may be controlled either by its relative error

$$\|s_k - \Delta x_k\| \leq \varepsilon_k \|s_k\|, \tag{1}$$

or its relative residual

$$\|r_k\| \leq \varepsilon_k \|F(x_k)\| \ . \tag{2}$$

Note that (1) requires an error estimate which may be hard to get, whereas the residual in (2) is in most cases cheaply available. Moreover, (1) requires a unique solution Δx of the linearized problem to be defined and thus does not apply to the underdetermined situation directly.

To make the invariance properties precise, let $\mathrm{Aff}(X)$ denote the ring of affine isomorphisms of X, i.e. maps of the form $Tx = Ax + b$ with some continuously invertible linear map $A = T' \in \mathrm{GL}(X)$ and $b = T(0) \in X$. Then, $\mathrm{Aff}(Y)$ and $\mathrm{Aff}(X)$ act on the maps $F : X \to Y$ by the covariant transformation $F \mapsto T_*F := T \circ F$ for $T \in \mathrm{Aff}(Y)$ and the contravatiant transformation $F \mapsto T^*F := F \circ T$ for $T \in \mathrm{Aff}(X)$, respectively. Obviously, the covariant transformation T_* does not alter the solution of the nonlinear problem $F(x) = 0$, since

$$F(x) = 0 \iff TF(x) = (T_*F)(x) = T(0) \ .$$

What is changed, is the solution's *characterization*. This is different from the contravariant transformation T^*. Since

$$F(x) = 0 \iff (T^*F)(T^{-1}x) = 0 \ ,$$

we do not change the nonlinear problem but the *representation* of its solution x.

Newton's method is (as successive linearization) invariant with respect to both transformations. More precisely, if we consider the covariant transformation T_* for $T \in \mathrm{Aff}(Y)$ and apply Newton's method to the transformed equation $T_*F(x) - T(0) = 0$, the Newton iterates x_k remain the same. On the other hand, the iterates transform in the same way as the solution does if we apply the contravariant transformation T^* for $T \in \mathrm{Aff}(X)$.

It is natural to ask for convergence theorems for Newton's method that respect at least one of these invariance properties. Concentrating on the covariant transformations of the image space, Deuflhard and Heindl [DH79] were led to so-called *affine invariant Lipschitz conditions* of the form

$$\|F'(y)^{-1}(F'(x+tv) - F'(x))v\| \le t\omega \|v\|^2$$

for all $x, y \in D$, $v = y - x$, and $t \in [0, 1]$. They only employ the norm in the domain X. In the context of an inexact Newton method (see [Deu91] and [Hoh93] for convergence theorems of that type), we encounter some problems. Most importantly, the affine covariant characterization naturally leads to an accuracy matching for the error as in (1) which, as already mentioned, is often hard to apply.

Accordingly, one may use an affine contravariant Lipschitz condition like

$$\|(F'(x+tv) - F'(x))v\| \le t\omega \|F'(x)v\|^2 \tag{3}$$

which only uses the norm in the image space Y. The following theorem states a convergence result based on this condition. It is a bit lengthy, because we include the undetermined case by restricting the corrections to suitable subsets $V(x)$ of X. Moreover, we introduce a parameter $0 \le \beta \le 1$ describing the "exactness" of the method in order to obtain a smooth transition from exact to inexact Newton methods. Setting $\beta = 0$ corresponds to the exact Newton iteration, while $\beta = 1$ distributes the error in equal parts on the Newton iteration and the inexact solution of the Newton equation.

THEOREM 1.1. *Let $F : D \subset X \to Y$ be a Gâteaux-differentiable mapping $D \subset X$ an open convex subset of X, and $\{V(x)\}_{x \in D}$ a family of subsets $V(x) \subset X$ with $V(x) \cap \ker F'(x) = \{0\}$. We assume that there is a constant $\omega > 0$ such that*

$$\|(F'(x+tv) - F'(x))v\| \le t\omega \|F'(x)v\|^2 \tag{4}$$

for all $x \in D$, $t \in [0, 1]$, and $v \in V(x)$ such that $x + v \in D$. Moreover, suppose that an inexact Newton sequence $\{x_k\}_{k=0,1,\ldots}$ exists in D such that for all $k \in \mathbb{N}$ and some $0 \le \beta \le 1$

a) the corrections $s_k = x_{k+1} - x_k$ are in $V(x_k)$,

b) the residuals $r_k = F(x_k) + F'(x_k)s_k$ *are bounded by*

$$\|r_k\| \le \varepsilon_k \, \|F'(x_k)s_k\| \quad \textit{where} \quad \varepsilon_k = \frac{\beta}{2} \min(1, \omega \, \|F'(x_k)s_k\|) \,, \tag{5}$$

c) the initial guess x_0 *satisfies*

$$\omega \, \|F(x_0)\| < h_{\max} := \frac{2-\beta}{1+\beta} \,. \tag{6}$$

Then the residuals $F(x_k)$ *converge quadratically to zero.*

The proof is simple and to be found in [Hoh93]. As for almost all convergence theorems for Newton-like methods, the whole key is the judicious application of the fundamental theorem of calculus. The whole formulation only involves the norm in the image space Y. That is why we can not prove convergence of the iterates x_k but only of the residuals $F(x_k)$.

If X is a Hilbert space and $\beta = 0$, i.e., in the exact case, we may substitute the orthogonal complement of the Jacobian's kernel for the restriction space:

$$V(x) := (\ker F'(x))^{\perp} \,.$$

This results in the Gauss Newton method. In the inexact case $\beta \neq 0$, this choice is not realistic, since s_k is in most cases only approximately orthogonal to the $\ker F'(x_k)$. Hence, the most suitable choice for $V(x)$ appears to be the 'algorithmic' restriction

$$V(x) := \begin{cases} \{s_k\} & \text{if } x = x_k \\ \emptyset & \text{if } x \notin \{x_k\}. \end{cases}$$

Introducing the *Kantorovitch quantities* $h_k := \omega \, \|F'(x_k)s_k\|$, the demand (5) on the inner residuals reads

$$\|r_k\| \le \varepsilon_k \, \|F'(x_k)s_k\| \,, \quad \text{where} \quad \varepsilon_k := \frac{\beta}{2} \, \min(1, h_k) \,. \tag{7}$$

This may be easily transferred into a matching strategy for the relative residuals

$$\|r_k\| \le \eta_k \, \|F(x_k)\|, \quad \text{where} \quad \eta_k := \frac{\varepsilon_k}{1+\varepsilon_k} = \varepsilon_k + O(\varepsilon_k) \,,$$

which obviously implies (7). This argument also holds in what follows: Up to $O(\varepsilon_k)$ the residual norm $\|F(x_k)\|$ and $\|F'(x_k)s_k\|$ are interchangeable.

Computational Estimates. To arrive at an implementable control mechanism, we have to replace the analytic quantities h_k with computational available estimates. Here, we proceed as in [Deu91]. From Theorem 1.1 we obtain the inequalities

$$\|F(x_{k+1})\| \le \frac{1+\beta}{2(1-\varepsilon_k)} h_k \, \|F(x_k)\| \tag{8}$$

and

$$h_{k+1} \le \frac{1+\beta}{2(1-\varepsilon_{k+1})} h_k^2 \,. \tag{9}$$

The first inequality (8) leads to the affine contravariant *a posteriori Kantorovitch estimate*

$$[h_k] := \frac{2(1-\varepsilon_k)}{1+\beta} \frac{\|F(x_{k+1})\|}{\|F(x_k)\|} \le h_k \,. \tag{10}$$

For the accuracy matching (7) we also need an estimate for h_k before the correction s_k is computed. Employing the quadratic convergence property, we derive the a priori estimate

$$[h_k] := \frac{1+\beta}{2(1-\varepsilon_{k-1})}[h_{k-1}]^2 . \tag{11}$$

Thus, we arrive at the affine contravariant matching strategy

$$[\varepsilon_k] := \frac{\beta}{2}\min(1,[h_k]) . \tag{12}$$

REMARK 1. It is rather easy in incorporate the inexact Gauss Newton iteration in the framework of a continuation method. Regarding the results by of Deuflhard, Fiedler and Kunkel [DFK87] from the residual oriented viewpoint, we arrive at almost the same stepsize control mechanism. For details, we refer to [Hoh93].

2. Application to BVPs of ODEs. We consider two point boundary value problems of ordinary differential equations of m-th order in $\mathbb{R}^n$

$$x^{(m)} = f(x,t) \;\text{ on } I=[a,b], \text{ and } r(x) = r(x(a),x(b)) = 0. \tag{13}$$

Here x not only denotes the function $x(t) \in \mathbb{R}^n$ but also the vector $x(t) \in \mathbb{R}^{mn}$ of the values of x and its derivatives $x', \ldots, x^{(m-1)}$ up to order $m-1$. Moreover, $f : \mathbb{R}^{mn} \times I \to \mathbb{R}^n$ and $r : \mathbb{R}^{mn} \times \mathbb{R}^{mn} \to Z = \mathbb{R}^{mn}$ are continuously differentiable mappings. The generalization to more general boundary conditions is straightforward.

To apply the inexact Newton method, we reformulate (13) as a nonlinear integral equation

$$F(x) := \begin{pmatrix} V(x) \\ r(x) \end{pmatrix} = 0 , \tag{14}$$

where $V(x)$ is the Volterra operator

$$V(x)(t) := x(t) - x(a) - (b-a)^{m-1}\int_a^t (t-s)^{m-1} f(x(s),s)\,ds .$$

If f and f' are uniformly bounded with respect to x by some polynomial, F is a well-defined Gâteaux-differentiable mapping

$$F : C^{m-1}(I) \longrightarrow L^2 \times Z .$$

Of particular interest are non degenerate boundary conditions, e.g.

$$r(x) = \beta x - z, \quad \beta x = B_1 x(a) + B_2 x(b),$$

where $B_1, B_2 \in \mathbb{R}^{mn\times mn}$ and the restriction of β on the space $\mathbb{P}_{m-1} = \ker D^m$ of polynomials of degree less than m induces an isomorphism

$$\beta : \mathbb{P}_{m-1} \xrightarrow{\cong} Z = \mathbb{R}^{mn}. \tag{15}$$

In this particular case, we can incorporate the boundary conditions in the integral equation and solver $F(x) = 0$, where $F : C^{m-1}(I) \longrightarrow L^2$ is the Fredholm operator

$$(Fx)(t) := x(t) - (\beta^{-1}z)(t) - \int_a^b G(t,s) f(x,s)\,ds \tag{16}$$

with Green's function $G(t,s)$ for the linear differential operator $L = D^m$ and the homogeneous boundary conditions $\beta x = 0$. In any case, the integral equation fulfills the formal requirements of theorem (1.1) for the inexact Newton method. Thus, we obtain a sequence of linear BVPs (again in their integral formulation) which have to be solved up to the residual prescribed by the surrounding inexact Newton iteration.

3. H-p Collocation for Linear BVPs. The two most successful approaches to the numerical solution of boundary value problems of ODEs seem to be the *local* approach by *multiple shooting* and the *global* one by *collocation*. Roughly speaking, multiple shooting is an efficient method for so-called *time-like* problems (the independent variable t has a distinguished direction) and strict accuracy demands, but needs good initial data (or lots of numerical insight) to converge at all. On the other hand, collocation (of fixed low order) is actually often used to produce these initial data, since the Newton method belonging to it converges much better. However, it becomes less efficient when it comes to strong accuracy demands. Moreover, *space-like* BVPs are mainly the realm of symmetric collocation methods.

The efficiency of shooting methods is mainly due to the adaptive order and stepsize control of the numerical integrators used to compute the flow and to solve the variational equation. In this section we try to combine the advantages of both methods and construct a collocation method with variable local orders. This line of thought has already attracted much interest in the PDE community (cf. [GB86], [DORH89], [ODRW89]), where the so-called h-p methods become more and more popular. The same idea also led to the very efficient numerical methods for countable systems of ODEs (so-called CODEs, cf. [Wul92] [DW94]).

The h-p collocation is based on local refinement by bisection and variable local orders. We preferred this standard approach from finite elements over regridding techniques, because we can store up local information due to the linearity of the BVP (in the context of the multilevel Newton method). Furthermore, the transition to BVPs in two or more dimensions seems to be much easier.

Since space is limited, we can not explain the algorithm in full detail, but try to sketch its most important ingredients. For more information we have to refer to the author's thesis [Hoh93]. We start with a short discription of the overall algorithm.

ALGORITHM 1.

1. *Compute the collocation solution on the present grid.*
2. *Estimate local and global residuals.*
3. *Check for convergence (overall residual estimate less than the required accuracy)*
4. *Choose for each subinterval the optimal refinement and order using an h-p error model.*
5. *Compute a threshold value which is to be the biggest local residual of the next level.*
6. *Apply the optimal refinement and order to all subintervals whose current residual estimate is bigger than the threshold.*

Steps 4 to 6 are responsible for the equidistribution of the local residual. We only refine a subinterval or increase its order, if the residual is still too big. Observe that due to the linearity of the BVP one may keep the local collocation matrices if the corresponding subinterval (inclusive its order) remains unaltered. Thus, we directly use information computed on the previous levels. Proceeding further, we give the notions mentioned above a more precise meaning.

Collocation on h-p grids. We have to introduce some notation. We define an *h-p grid* $\Delta = (\{t_i\}, \{p_i\})$ as a partition

$$a = t_1 < t_2 < \cdots < t_{N+1} = b$$

of the basic interval $[a, b]$ into N subintervals $J_i = [t_i, t_{i+1}]$ of length $h_i = t_{i+1} - t_i$ and local orders p_i for $1 \le i \le N$. By $\Delta + k$ we denote the grid

$$\Delta + k := (\{t_i\}, \{p_i + k\})$$

obtained by adding the integer k to the local orders. $\mathbb{P}_{\Delta+k}$ will denote the space of piecewise polynomials of local degrees $p_i + k$, i.e.,

$$x \in \mathbb{P}_{\Delta+k} \iff x|_{J_i} \in \mathbb{P}_{p_i+k} \text{ for all } i = 1, \ldots, N.$$

Collocation means that we are looking for a spline $x \in S_{\Delta,m} := \mathbb{P}_{\Delta+m-1} \cap C^{m-1}[a,b]$ which satisfies boundary conditions and the ODE at all collocation points. More explicitly, we have to compute polynomials $x_i \in \mathbb{P}_{p_i+m-1}$ on J_i such that

a) the overall solution is in $C^{m-1}[a,b]$, i.e.,

$$x_{i-1}(t_i) = x_i(t_i) \in \mathbb{R}^{mn} \quad \text{for } i = 2, \ldots, N$$

b) the local collocation conditions

$$x_i^{(m)}(t) = f(x_i(t), t)$$

are satisfied for $t = t_i + c_j h_i$, $1 \le i \le N$ and $1 \le j \le p_i$,

c) the boundary condition is fulfilled, i.e.,

$$r(x) = 0 \, .$$

Here, the $c_j \in [0,1]$ are the (typically Gaussian) collocation nodes.

Residual Estimation. In the framework of the inexact Newton method, we have to be able to compute the norms $\|F(x_k)\|$ and $\|r_k\|$ of the outer and inner residuals, respectively. In our present context this means to estimate the L^2-norm of the Volterra (14) or Fredholm (16) operator applied to the piecewise polynomial approximation $x \in S_{\Delta,m}$. To this end, we approximate the residual function $F(x)$ by a piecewise polynomial in $S_{\Delta,m+k}$, $k > 0$, and compute its L^2-norm by an appropriate quadrature formula. For details, see [Hoh93]. The same technique is also used to provide estimates for the local residuals on each subinterval J_i.

Error model. Given a subinterval J of length $\bar{h}$ and order $\bar{p}$, we would like to know the local residual obtained for different stepsizes h and orders p. To this end, we construct a *local h-p error model* $\varepsilon(h,p)$ that depends on three parameters to be estimated in the algorithm. By standard collocation theory we know that the error of the collocation solution with respect to the norm $\|\cdot\|_\infty$ is $O(h^p)$ for sufficiently smooth solutions. By continuity, the same estimate holds for the residual measured in the L^2-norm, i.e.,

$$\|Fx\| \le Ch^p \, ,$$

where F is the Volterra or Fredholm operator as above. Of course, the constant C contains bounds for the higher derivatives depending on p so that the error model $\varepsilon(h,p) = Ch^p$ with the single parameter C is not realistic. The standard choice for fixed order methods (e.g. linear finite elements) is

$$\varepsilon(h) = Ch^\gamma$$

including a second parameter $\gamma > 0$. Combining this approach with a third term describing the variable order, we are lead to the h-p error model

$$\varepsilon(h,p) = Ch^\gamma \alpha^p \tag{17}$$

depending on the three parameters $C, \alpha, \gamma \ge 0$. The *stepsize coefficient* γ characterizes the influence of refinement while the *order coefficient* α is responsible for order variations. Once we know the coefficients α and γ, we obtain the desired estimate for the local residuals by

$$\varepsilon(h,p) = \varepsilon(\bar{h},\bar{p}) \left(\frac{h}{\bar{h}}\right)^\gamma \alpha^{p-\bar{p}} \, .$$

To compute the order coefficient α, we use the residual estimate for the order $p = \bar{p} - 1$ leading to

$$\alpha = \sqrt[p-\bar{p}]{\frac{\varepsilon(\bar{h},p)}{\varepsilon(\bar{h},\bar{p})}} = \frac{\varepsilon(\bar{h},\bar{p})}{\varepsilon(\bar{h},\bar{p}-1)} \ .$$

Using this estimate for α, we may employ any other residual estimate to compute γ by

$$\gamma = \log_{h/\bar{h}} \left(\frac{\varepsilon(h,p)}{\varepsilon(\bar{h},\bar{p})} \alpha^{\bar{p}-p} \right) \ .$$

Optimal order and refinement. Now we have the main tools at hand to choose an optimal order and refinement for the next level. Since the error model is only feasible in a neighbourhood of the current stepsize and order $(\bar{h},\bar{p})$, we only consider pairs (h,p) from a so-called *order-stepsize window*

$$W(\bar{h},\bar{p}) := \{(\bar{h},\bar{p}+1\} \cup \left\{(\bar{h},p) \mid \ \bar{p}/2+1 \le p \le \bar{p}\right\} \ .$$

In other words, we either increase the order by one or refine the interval and choose a new order $\bar{p}/2+1 \le p \le \bar{p}$. We call an h-p pair $(h,p) \in W(\bar{h},\bar{p})$ *optimal*, if it minimizes the *work per accuracy* measured by the amount of work $\mathcal{A}(p)$ times the error model (as the expected residual), i.e.,

$$\varepsilon(h,p) \cdot \mathcal{A}(p) \cdot \frac{\bar{h}}{h} \ = \ \min! \tag{18}$$

Here, we have to take into account that the local amount of work doubles if we subdivide a subinterval. Moreover, we add the effort for the elimination of the local boundary values in that case. Nonetheless, we neglect the increased effort necessary to solve the rest of the *global* system, if more subintervals are present. This is part of our quite conservative strategy to use high orders only if they really pay off.

Multilevel Newton Method. The last sections give us a black box solver for the solution of a linear BVP up to a prescribed residual. The result is a collocation solution on an adaptively chosen h-p grid. As the final task, we only have to fit this linear solver into the framework of the inexact Newton method arriving at the adaptive h-p collocation for nonlinear BVPs.

4. Examples. In this section we give some examples for the new multilevel Newton h-p collocation algorithm, illustrating the h-p collocation for linear problems and the multilevel Newton method for nonlinear problems based on the Fredholm (Example 1) and the Volterra formulation (Example 2).

EXAMPLE 1. *Nonlinear transition layer.* This example is a well-known test problem taken from [AMR88] (Example 9.8). We consider the scalar nonlinear BVP given by

$$\varepsilon x'' + x x' - x = 0 \ \text{ on } [0,1] \ \text{ and } \ x(0) = x(1) = \frac{1}{2} . \tag{19}$$

Here, the solution has for small $0 < \varepsilon \ll 1$ a rapid transition layer at $t = 0$. Figure 1 presents the solution process for $\varepsilon = 10^{-3}$ and a required residual of $\text{tol} = 10^{-6}$. We display the iterates of the multilevel Newton method together with the corresponding h-p grids, i.e., the successively finer grids obtained by adding the inexact Newton corrections. The left column shows the approximate collocation solution on the indicated level while the right column displays the corresponding h-p grid. Here, we plotted the local orders versus the midpoints of the intervals. Figure 2 displays the corresponding Newton corrections obtained as the h-p collocation solutions of the linear subproblems. Here we see that the grids for the linear subproblems may be much coarser then the resulting grid of the nonlinear solution (obtained as the union of the correction grids). Hence, the linear systems are much easier to solve than the linear problems obtained by the standard approach 'linearization after discretization'.

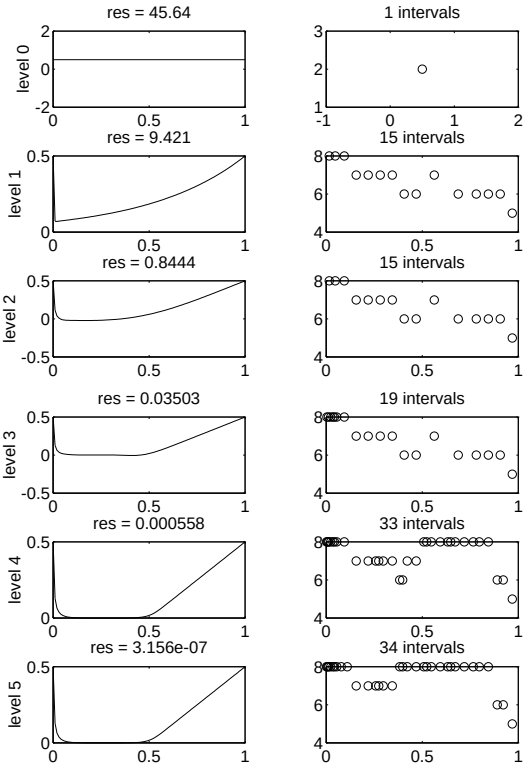

FIG. 1. *Adaptively chosen grids for (19) with* $\varepsilon = 10^{-3}$ *and* $\mathrm{tol} = 10^{-6}$

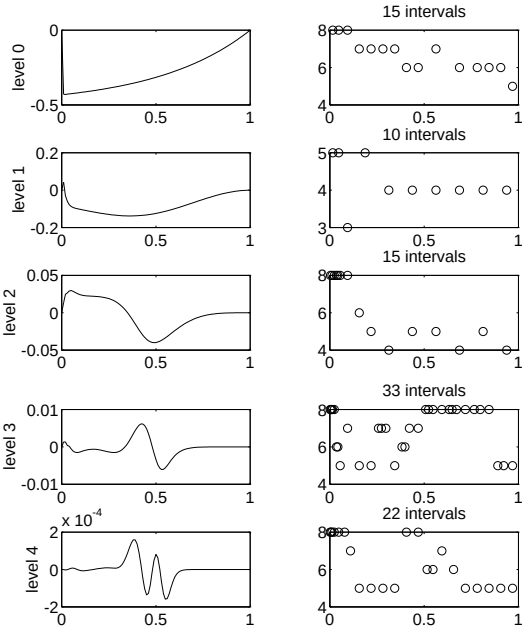

FIG. 2. *Newton corrections for (19) with* $\varepsilon = 10^{-3}$ *and* $\mathrm{tol} = 10^{-6}$

EXAMPLE 2. *Chemical Oscillator.* As a second example consider the following system of five autonomous differential equations modelling a chemical oscillator (cf. [See81]).

$$x_1' \quad = \quad j - k_1 x_1 - k_4 x_1 x_4 + k_{-4}(E_{\mathrm{tot}} - x_4 - x_5)$$

$$
\begin{aligned}
x_2' &= k_1 x_1 - k_2 x_2 \\
x_3' &= k_2 x_2 - k_3 x_3 - k_5 x_3 (E_{\rm tot} - x_4 - x_5) + (k_{-5} + k_6) x_5 \\
x_4' &= -k_4 x_1 x_4 + k_{-4} (E_{\rm tot} - x_4 - x_5) + k_6 x_5 \\
x_5' &= k_5 x_3 (E_{\rm tot} - x_4 - x_5) - (k_{-5} + k_6) x_5
\end{aligned} \tag{20}
$$

We compute a periodic solution for the parameters $E_{\rm tot} = 1$, $k_1 = k_2 = k_3 = 1$, $k_4 = k_{-5} = 2000$, $j = 100$, $k_{-4} = k_5 = 100$, $k_6 = 600$. Figure 3 shows the four Newton iterates necessary for a required

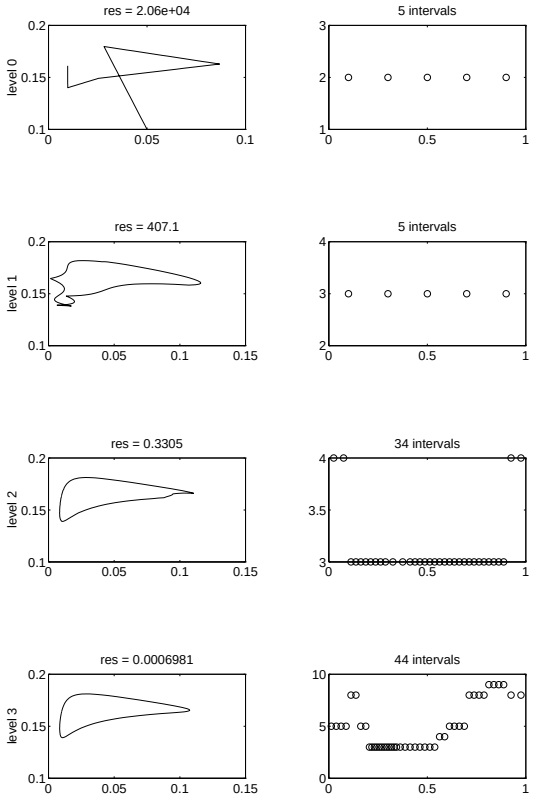

FIG. 3. *Newton iterates for the chemical oscillator*

accuracy tol $= 10^{-3}$. Here, we plotted x_5 versus x_4 in the left column. As initial guess we took the linear interpolant at 5 points which we obtained by an integration from $t = 0$ to $t = 3$ using the extrapolation code EULSIM. The initial value was $x = (9, 7, 5, 0.05, 0.1)$ and as initial guess for the period we set $T = 3$. The first picture clearly shows that this initial guess is far away from the periodic solution.

REMARK 2. Results combing an adaptive continuation method with the new BVP solver may be found in [Hoh93] and [Hoh94]. There, we compute a branch of periodic solution emanating at a Hopf bifurcation for a realistic model of a railway bogie moving on a straight track.

[ACR79] U. Ascher, J. Christiansen, and R. D. Russell. A collocation solver for mixed order systems of boundary value problems. *Math. Comput.*, 33:659–679, 1979.

[AMR88] U. M. Ascher, R. M. M. Mattheij, and R. D. Russell. *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*. Prentice Hall, Englewood Cliffs, New Jersey, 1988.

[Bad88] G. Bader. Solution of boundary value problems by collocation methods. Habilitationsschrift, Universität Heidelberg, 1988.

[Bor92] F. A. Bornemann. An adaptive multilevel approach to parabolic equations III. *Impact of Computing in Science and Engineering*, 4:1–45, 1992.

[BR81] R. E. Bank and D. J. Rose. Global approximate newton methods. *Numer. Math.*, 37:279–295, 1981.

[DES82] R. S. Dembo, S. C. Eisenstat, and T. Steihaug. Inexact newton methods. *SIAM J. Numer. Anal.*, 19:400–408, 1982.

[Deu91] P. Deuflhard. Global Inexact Newton Methods for Very Large Scale Nonlinear Problems. *Impact of Computing in Science and Engineering*, 3(4):366–393, 1991.

[DFK87] P. Deuflhard, B. Fiedler, and P. Kunkel. Efficient Numerical Pathfollowing beyond Critical Points. *SIAM J. Numer. Anal.*, 18:949–987, 1987.

[DH79] P. Deuflhard and G. Heindl. Affine invariant convergence theorems for newton's method and extensions to related methods. *SIAM J. Numer. Anal.*, 16:1–10, 1979.

[DORH89] L. Demkowicz, J. T. Oden, W. Rachowicz, and O. Hardy. Toward a universal h-p adaptive finite element strategy, Part 1. Constrained approximation and data structure. *Comput. Methods Appl. Mech. Engrg.*, 77:79–112, 1989.

[DW94] P. Deuflhard and M. Wulkow. Simulationsverfahren für die polymerchemie. Preprint SC 94–22, Konrad-Zuse-Zentrum, Berlin, 1994.

[GB86] W. Gui and I. Babuška. The h, p and h-p versions of the finite element method in one dimension, Part 1 to 3. *Numer. Math.*, 49:577–683, 1986.

[Hoh93] A. Hohmann. *Inexact Gauss Newton Methods for Parameter Dependent Nonlinear Problems*. PhD thesis, Freie Universität Berlin, 1993.

[Hoh94] A. Hohmann. Object Oriented Design of Multilevel Newton and Continuation Methods. In Allan Vermeulen, editor, *Proceedings of the Second Annual Object-Oriented Numerics Conference OON-SKI '94, Sunriver*. Rogue Wave Software, Corvallis, Oregon, 1994.

[LW92] J. Lang and A. Walter. A Finite Element Method Adaptive in Space and Time for Nonlinear Reaction-Diffusion Systems. *Impact of Computing in Science and Engineering*, 4:269–314, 1992.

[ODRW89] J. T. Oden, L. Demkowizc, W. Rachowicz, and T. A. Westermann. Toward a universal h-p adaptive finite element strategy, Part 2. A posteriori error estimation. *Comput. Methods Appl. Mech. Engrg.*, 77:113–180, 1989.

[Rhe86] W. C. Rheinboldt. *Numerical analysis of parametrized nonlinear equations*. John Wiley and Sons, New York, 1986.

[See81] F. F. Seelig. Unrestricted Harmonic Balance II. Application to Stiff ODE's in Exzyme Catalysis. *J. of Math. Biology*, 12:187–198, 1981.

[Wul92] M. Wulkow. Adaptive Treatment of Polyreactions in Weighted Sequence Spaces. *Impact of Computing in Science and Engineering*, 4(2):153–193, 1992.

[Ypm84] T. J. Ypma. Local convergence of inexact newton methods. *SIAM J. Numer. Anal.*, 21:583–590, 1984.