Applications of Mathematics

Erich Bohl; Johannes Schropp

Discrete evolutions: Convergence and applications

Applications of Mathematics, Vol. 38 (1993), No. 4-5, 266–280

Persistent URL: http://dml.cz/dmlcz/104555

38 (1993) APPLICATIONS OF MATHEMATICS No. 4–5, 266–280

DISCRETE EVOLUTIONS: CONVERGENCE AND APPLICATIONS

ERICH BOHL, JOHANNES SCHROPP, Konstanz

Summary. We prove a convergence result for a time discrete process of the form

$$x(t+h) - x(t) = hV(h, x(t+\alpha_1(t)h), \ldots, x(t+\alpha_L(t)h))$$

$$t = T + jh,\ j = 0, \ldots, \sigma(h) - 1$$

under weak conditions on the function V. This result is a slight generalization of the convergence result given in [5]. Furthermore, we discuss applications to minimizing problems, boundary value problems and systems of nonlinear equations.

Keywords: discrete processes, continuous processes, convergence of discretisations, boundary value problems, minimizing problems, Newton's iteration and Newton's flow

AMS classification: 65L99

0. INTRODUCTION

0.1. The literature on discrete processes is full of iterations [4, 6, 12, 17] of the form

$$x(t+h) - x(t) = hV(x(t)). \tag{1}$$

This equation simply says that the new information $x(t+h)$ at the new time $t+h$ equals the old information $x(t)$ at the old time t plus a term which describes the variation occurring during the elapsed time h. This term is proportional to the time h with a function of proportionality which typically depends on the old information $x(t)$ which is an n-dimensional vector with real components.

In numerical analysis a standard problem is to solve a system $F(x) = 0$ of N equations with the same number of unknowns. A typical approach [2, 8, 14] is an iterative process

$$x(t+h) - x(t) = -hA(x(t))F(x(t)) \tag{2}$$

with a 'damping factor' h. $A(x(t))$ is a real (N, N)-matrix which typically tries to approximate the inverse of the Jacobian $DF(x(t))$. (2) is then sometimes called a damped Newton-like method.

Another standard problem [8, 9, 16] is the minimization of a real-valued function f of N unknowns:

$$f(x) \leqslant f(y) \quad \text{for} \quad y \in \mathbf{R}^N. \tag{3}$$

The widely used descent methods are of the general form

$$x(t+h) - x(t) = -hp(x(t)) \tag{4}$$

where the vector $p(x(t))$ sometimes tries to approximate the gradient grad $f(x(t))$ of the function f at the point $x(t)$.

Finally, multistep methods [7, 10, 11] for the numerical treatment of ordinary differential equations read

$$x(t+h) - x(t) = hV(h, x(t+h), x(t), \ldots, x(t-kh)) \tag{5}$$

where k is a fixed natural number and $h > 0$ denotes the step width. There are other types of multistep methods. However, most of them can be cast in the form (5).

Obviously, (1), (2), (4) and (5) are special cases of the process

$$x(t+h) - x(t) = hV(h, x(t+\alpha_1(t)h), x(t+\alpha_2(t)h), \ldots, x(t+\alpha_L(t)h)). \tag{6}$$

With the definition

$$z(t,h,x) := (x(t+\alpha_1(t)h), \ldots, x(t+\alpha_L(t)h)), \alpha_j(t) \in \mathbf{Z},\ j = 1, \ldots, L \tag{7}$$

(6) can be written as

$$x(t+h) - x(t) = hV(h, z(t,h,x)). \tag{8}$$

Depending on the integers $\alpha_j(t)$, (8) may be an explicit or implicit process. It is even possible that a successive solution as suggested by (6) is not possible.

0.2. In [5] we have proved a general convergence result on processes (6) under special conditions for the constants $\alpha_j, j = 1, \ldots, L$. In this paper we generalize this result slightly allowing for more freedom in choosing the integers α_j. Our theorem states, roughly speaking, that the solutions of (6) behave for $h > 0$ small enough like the solutions of the differential equation

$$\dot{x}(t) = V(0, z(t,0,x)) = V(0, x(t), \ldots, x(t)). \tag{9}$$

This has implications for the special processes (2) and (4) which might break down on the grounds that the 'damping factor' h is forced to be chosen unrealistically

small. This suggests that the first order system (9) is stiff and calls for a treatment by an A-stable method which typically is not explicit [7, 10, 11] as (2) or (4) but an implicit procedure. Our point is that an 'explicit' method like (2) to find a zero of an N-dimensional system $F(x) = 0$ or like (4) to treat the unconstrained minimization problem (3) might be inappropriate. This point was already made 20 years ago by Boggs [1]. His work seems to have been overlooked since then. We discuss this in more detail in Section 2 and give numerical examples for our point. In Section 1 we state the theorem and comment on the proof as far as the arguments change in contrast to our paper [5]. Hence, we concentrate on two parts of the proof where slight changes compared to [5] are necessary. Finally, in Section 3, we comment on two-point-boundary-value problems for second order scalar equations. Here, a convergence proof of the classical finite difference method is derived from our main theorem in Section 1.

We finally mention a technical point: We refer to formula (i) in Section j as (j.i) and we drop the Section index j if we refer to formula (i) in the current section.

1. The main result

1.1. We consider two real numbers $h > 0$, T defining the grid points $T + jh$, $j \in Z$. Furthermore, let $S > T$ define the natural number $\sigma(h)$ via

$$-h < S - (T + \sigma(h)h) \leqslant 0. \tag{1}$$

Finally, we consider integers

$$\alpha_k(T + jh) \in \mathbf{Z},\ |\alpha_k(T + jh)| \leqslant \varrho,\ \varrho \text{ independent of } j, h, k, \tag{2}$$

$$j = 0, \ldots, \sigma(h) - 1,\ k = 1, \ldots, L$$

to define $z(t, h, x)$ in (0.7) so that our system reads

$$x(t + h) - x(t) = hV(h, z(t, h, x))$$

$$t = T + jh,\ j = 0, \ldots, \sigma(h) - 1 \tag{3}$$

$$x(t + \alpha_k(t)h) = x(T) \text{ if } t + \alpha_k(t)h < T$$

$$x(t + \alpha_k(t)h) = x(T + \sigma(h)h) \text{ if } T + \sigma(h)h < t + \alpha_k(t)h.$$

We assume

$$V \in C(\mathbf{R}_+ \times \mathbf{R}^{LN}, \mathbf{R}^N) \tag{4}$$

$$\|V(h,u)\|_\delta \leqslant M \quad \text{if} \quad 0 \leqslant h \leqslant h_0,\ u \in \mathbf{R}^{LN}$$

where $\|\ \|_\delta$ denotes the maximum norm in $\mathbf{R}^N$. Then, (3) is a system of $N\sigma(h)$ equations of $N(\sigma(h)+1)$ unknowns: $\sigma(h)+1$ counts the number of grid points in the interval $[T, T+\sigma(h)h]$. Hence, there are N more unknowns than we have equations in (3). We might interpret the first line in (3) as a discrete evolutionary process which is observed in a finite time interval $[T,S]$. Any solution $x^h(t)$ of (3) is, in the first place, only defined at the grid points. However, we always consider $x^h(t)$ as a continuous function on the whole real axis $\mathbf{R}$ after linear interpolation between consecutive grid points and constant extension to the right of the most right grid point and to the left of the most left grid point. We refer to the resulting object as a grid function on $\mathbf{R}$.

1.2. Theorem. *Consider the system (3) under the assumptions (2) and (4). Let $h_n > 0$ be a real sequence tending to zero such that the corresponding system (3) has a solution $x^{h_n}(t)$ with*

$$\|x^{h_n}(T)\| = O(1) \quad \text{as} \quad h_n \to 0. \tag{5}$$

Then there is a function $v \in C^1[T,S]$ which solves the differential equation

$$\dot{v}(t) = V(0, z(t,0,v)),\ T \leqslant t \leqslant S. \tag{6}$$

Furthermore, there exists a subsequence of $x^{h_n}(t)$ such that

$$\max\{\|v(t) - x^{h_n}(t)\|_\delta : t \in [T,S]\} \to 0 \quad \text{as} \quad n \to \infty \tag{7}$$

holds. Notice that we did not change notation passing from the sequence $x^{h_n}(t)$ to a subsequence.

1.3. A proof of Theorem 1 is given in [5] for the special case

$$\alpha_k(T+jh) = -(k-2),\ k = 1,\dots,L,\ j = 1,\dots,\sigma(h),\ L \geqslant 2 \tag{8}$$

of the integers (2). Obviously, the uniform bound ϱ required in (2) is $\varrho = \max\{|k-2|: k = 1,\dots,L\}$ for (8). Assuming (8) we have

$$z(t,h,x) = (x(t+h), x(t), x(t-h), \dots, x(t-(L-2)h)). \tag{9}$$

We have also discussed in [5] that the assumption in the second line of (4) which requires a uniform and global bound for the function V is essentially without loss of generality: This is because we restrict our attention to a compact interval $[T,S]$

so that any solution of the differential equation (6) is bounded and we can replace $V(h,u)$ by the globally bounded function

$$V^{(n)}(h,u) = V(h, u\Phi(u/n))$$

with a natural number n appropriately chosen. Here, $\Phi(u)$ is defined via

$$\Phi(u) = \varphi(u_1)\varphi(u_2)\dots\varphi(u_{LN}),\ u = (u_1,\dots,u_{LN}),$$

$$\varphi \in C(\mathbf{R}),\ |\varphi(s)| \leqslant 1 \text{ in } \mathbf{R},$$

$$\varphi(s) = 1 \text{ if } |s| \leqslant 1,\ \varphi(s) = 0 \text{ if } |s| \geqslant 2.$$

We finally note that $z(t,0,v) = (v(t), v(t), \dots, v(t)) \in \mathbf{R}^{NL}$ follows from (0.7) so that the limiting differential equation (6) actually reads

$$\dot{v}(t) = V(0, v(t), \dots, v(t)),\ T \leqslant t \leqslant S. \tag{10}$$

1.4. We already mentioned in 0.1 that the processes (0.1), (0.2), (0.4) and (0.5) are special cases of (3) if we complete (3) by initial conditions. Obviously, then, the assumption (5) of our theorem is satisfied and the limiting differential equation (6) can be completed by the initial conditions

$$x^h(T) = w \in \mathbf{R}^N. \tag{11}$$

Then we have proved in [5] that convergence in the sense of

$$\max\{\|v(t) - x^h(t)\|_\delta : t \in [T,S]\} \to 0 \text{ as } h \to 0 \tag{12}$$

holds if the initial value problem

$$\dot{v}(t) = V(0, z(t,0,v)),\ T \leqslant t \leqslant S,\ v(T) = w \tag{13}$$

has at most one solution. Since this is generically the case, we have (12) in a generic situation. Notice, however, that we must assume the solvability of the system (3) along with the initial conditions (11) for any h appearing in (12).

In this sense, the special processes considered in 0.1 have an initial value problem as a continuation if the 'damping' h tends to zero. In particular, the differential equations to (0.1), (0.2), (0.4) and (0.5) are

$$\dot{x}(t) = V(x(t)),\ T \leqslant t \leqslant S,$$

$$\dot{x}(t) = -A(x(t))F(x(t)),\ T \leqslant t \leqslant S, \tag{14}$$

$$\dot{x}(t) = -p(x(t)),\ T \leqslant t \leqslant S,$$

$$\dot{x}(t) = V(0, x(t), \ldots, x(t)),\ T \leqslant t \leqslant S,$$

respectively. If, for example, (0.2) is exactly Newton's method to find a zero for F, then the second differential equation in (14) is Newton's flow. Similarly, if we use in (0.4) the gradient direction to find a minimum of f, then the continuous extension of (14) is the differential equation

$$\dot{x}(t) = -\operatorname{grad} f(x(t)),\ T \leqslant t \leqslant S. \tag{15}$$

1.5. The proof of Theorem 1.2 is only slightly different from the proof given in [5] so that we give only some hints to the arguments which are new here. Therefore, we concentrate on two parts in the proof where the difference appears in the more general situation. First of all, linear interpolation between grid points yields the representation

$$x^h(s) - x^h(t) = V(h, z(t, h, x^h))(s - t) \tag{16}$$

$$\text{for } t \leqslant s \leqslant t + h,\ t = T + jh,\ j = 0, \ldots, \sigma(h) - 1$$

from which we conclude

$$\|x^h(s) - x^h(t)\|_\delta \leqslant M|s - t|, \text{ if } T \leqslant s,\ t \leqslant S \tag{17}$$

as in [5]. A consequence is

$$\|x^h(t)\|_\delta \leqslant \|x^h(T)\|_\delta + \|x^h(T) - x^h(t)\|_\delta$$

$$\leqslant M|t - T| + \|x^h(T)\|_\delta \leqslant M|S - T| + \|x^h(T)\|_\delta \tag{18}$$

$$\text{if } T \leqslant t \leqslant S$$

so that our assumption (5) immediately implies

$$\max\{\|x^h(t)\|_\delta : T \leqslant t \leqslant S\} = O(1) \text{ as } h \to 0. \tag{19}$$

The Lipschitz condition (17) uniform in $h > 0$ and the bound (19) uniform in $t \in [T, S]$ and in $h > 0$ allows the application of the Arzèla-Ascoli-Theorem which implies that there exists a function $v(t) = (v_1(t), \ldots, v_N(t))$, $T \leqslant t \leqslant S$ with continuous components $v_j \in C[T, S], j = 1, \ldots, n$ satisfying

$$\max\{\|x^{h_n}(t) - v(t)\|_\delta : T \leqslant t \leqslant S\} \to 0 \text{ as } n \to \infty. \tag{20}$$

In an attempt to prove that the function $v(t)$ has actually C^1-components, we must estimate the difference

$$\int_T^t V(0, z(s,0,v))ds - \sum_{r=0}^{j} hV(h, z(t_r, h, x^h))$$

$$= \int_T^t V(0, z(s,0,v))ds - \sum_{r=0}^{j} hV(0, z(t_r, 0, v)) \tag{21}$$

$$+ \sum_{r=0}^{j} h[V(0, z(t_r, 0, v)) - V(h, z(t_r, h, x^h))]$$

(compare [5]). Here, $t \in (T, S], h > 0$ are fixed numbers and $0 \leqslant j := j(h,t) < \sigma(h)$ is uniquely defined by

$$t_j \leqslant t < t_{j+1}, t_j := T + jh. \tag{22}$$

In particular, we must show that

$$\max\{\|V(0, z(s,0,v)) - V(h, z(s,h,x^h))\|_\delta : s = T + kh,\ k = 0, \ldots, \sigma(h) - 1\} \tag{23}$$

$$\to 0 \quad \text{as} \quad h \to 0$$

(compare [5]). To this end, we consider the two z-components in (23)

$$z(s,0,v) = (v(s), \ldots, v(s)) \tag{24}$$

$$z(s,h,x^h) = (x^h(s + \alpha_1(s)h), \ldots, x^h(s + \alpha_L(s)h)).$$

The individual vector components can be estimated as follows

$$\|v(s) - x^h(s + \alpha_k(s)h)\|_\delta \leqslant \|v(s) - x^h(s)\|_\delta + \|x^h(s) - x^h(s + \alpha_k(s)h)\|_\delta$$

$$\leqslant \max\{\|v(s) - x^h(s)\|_\delta : T \leqslant s \leqslant S\} + M|\alpha_k(s)|h \tag{25}$$

$$\leqslant \max\{\|v(s) - x^h(s)\|_\delta : T \leqslant s \leqslant S\} + M\varrho h.$$

This shows that they are as close together as we want if we choose $h > 0$ appropriately small. Moreover, the estimate in (25) is uniform in h. Therefore, if h is small enough, all components of the independent variables of the V-function appearing in (23) are uniformly close together. By the uniform continuity of the continuous function V on a compact set, the limit (23) follows.

The rest of the proof is verbatim the same as in the special case treated in [5].

2. Newton's method, gradient method and conjugate gradient method.

2.1. In this Section we consider more closely the process (0.2) in the special form of damped Newton's method

$$x(t+h) - x(t) = -hDF(x(t))^{-1}F(x(t)),\ x(0) = u. \tag{1}$$

The corresponding continuous extension is the initial value problem

$$\dot{x}(t) = -DF(x(t))^{-1}F(x(t)),\ x(0) = u \tag{2}$$

whose solutions $\varphi(t, u)$ are known as Newton's flow.

It is obvious that any stationary solution $\bar{x}$ of (2) such that the Jacobian $DF(\bar{x})$ is non-singular, satisfies the condition

$$F(\bar{x}) = 0. \tag{3}$$

The derivative of the right-hand side of (2) at such a stationary point is

$$D(-DF(x)^{-1}F(x))_{x=\bar{x}} = -I. \tag{4}$$

This means that the stationary point is (asymptotically) stable, implying that there is a neighbourhood $U(\bar{x})$ such that for any initial condition $x(0) = u \in U(\bar{x})$ the solution $\varphi(t, u)$ tends to $\bar{x}$ as $t \to +\infty$. This takes care of the local behaviour of $\varphi(t, u)$ in a neighbourhood of a stationary point. A global result reads as follows and will be proven in a forthcoming paper:

Theorem [15]. *Let $\varphi(t, u)$ be a solution of (2) for all $t \geqslant 0$. If its ω-limit set is nonempty, then there exists a $z \in \mathbf{R}^N$ satisfying*

$$F(z) = 0,\ \varphi(t, u) \to z \ \text{ as } \ t \to +\infty.$$

We may expect that damped Newton's method (1) follows Newton's flow closely if $h > 0$ is small enough. Therefore, the theorem and the local argument given above describe all initial points such that damped Newton's method (1) is convergent to a zero of the vector field F. Hence, if Newton's method ($h = 1$) in (1) is not convergent, the damping technique can only provide a remidy if the initial approximation $x(0) = u$ is chosen as just described. In this sense, all basins of attraction for (1) are essentially given by the basins of attraction of Newton's flow (2).

Even if the damping technique is successful and leads theoretically to a convergent sequence, it is well possible that the damping h may have to be taken unacceptably small so that practically (1) is not a reasonable method to solve (3). This situation will occur if the initial value problem (2) happens to be a stiff problem. Then, typically, explicit methods as (1) are not reasonable. To solve (2), we rather must go for A-stable formulae which are typically implicit. We remark that, to our knowledge, this was done by Boggs [1] for the first time.

2.2. This argument can be demonstrated very nicely for the process (0.4) if we use the gradient method

$$x(t+h) - x(t) = -h \operatorname{grad} f(x(t)),\ x(0) = u \tag{5}$$

as a special case. The corresponding continuous extension is

$$\dot{x}(t) = -\operatorname{grad} f(x(t)),\ x(0) = u \tag{6}$$

(compare (1.14)). If (6) turns out to be stiff, then the explicit scheme (5) to solve (6) needs an unrealistically small step width h so that the gradient method (5) practically cannot be used. It must be replaced by a stiff (A-stable) solver. This situation occurrs quite often in the applications. A particularly simple case is the following:

2.3. In the theory of enzyme actions, Michaelis-Menten kinetics is very popular. It is the simplest possibility for a mechanism governing an enzyme which transforms a substrate into a product. The Michaelis-Menten theory says that the velocity v to transform the concentration w of a substrate depends on w via the rational expression

$$v = g(w, \mu, K) = \frac{\mu w}{K + w},\ \mu > 0,\ K > 0 \tag{7}$$

with appropriate constants μ, K. This expression looks very simple. It, however, leads to a very uncomfortable fitting problem since the constant K typically is positive but much smaller than μ. Then, $\frac{\partial g}{\partial w}(w, \mu, K)$ explodes if w is in a neighbourhood of zero.

w	.3330	.1670	.0833	.0416	.0208	.0104	.0052
v	3.636	3.636	3.236	2.666	2.114	1.466	.866

Table 1: Invertase reaction (taken from [3, 13])

To be more precise, consider the pairs (w, v) given in Table 1, which are taken from actual measurements in [13]. These numbers belong to the rational expression (7) with appropriate constants μ, K. To recover these constants from the data in Table 1, we apply the method of least squares and minimize the sum of squares

$$f(x) = f(\mu, K) = \sum_{i=1}^{7} (g(w_i, \mu, K) - v_i)^2. \tag{8}$$

Using (5) to solve the least square problem, we find the solution

$$\bar{x} = (\bar{\mu}, \bar{K}) = (3.91, 1.788E-2)$$

for sufficiently many starting points (μ_0, K_0) and $h = 10^{-5}$. There is no way of obtaining convergence for a step-width h which is greater than 10^{-5} so that the computational work is tremendous.

To demonstrate the dependence of the step-size h used in the process (5) on the stiffness of the underlying differential equation (6), we modify the function g in (7) and consider

$$\tilde{g}(w, \alpha, \nu, \kappa) = \frac{\nu w}{\alpha\kappa + w}. \tag{9}$$

The corresponding sum of squares reads

$$\tilde{f}(\alpha, \nu, \kappa) = \sum_{i=1}^{7} (\tilde{g}(w_i, \alpha, \nu, \kappa) - v_i)^2. \tag{10}$$

We fix α and minimize (10) with respect to ν, κ. We can regulate the stiffness of the underlying differential equation

$$\dot{x} = -\operatorname{grad} \tilde{f}(\alpha, x), \quad x := (\nu, \kappa) \tag{11}$$

varying the constant α. To demonstrate this, we give the eigenvalues of the linearization of the right-hand side in (11) for six values of α in Table 2.

α	0.01	0.1	1	10	100	1000
λ_1	0.762	2.85	2.9	2.9	2.9	2.9
λ_2	8.074	215.7	$2.119 \cdot 10^4$	$2.118 \cdot 10^6$	$2.117 \cdot 10^8$	$2.1 \cdot 10^{10}$
stiffness: $\frac{\lambda_2}{\lambda_1}$	10.59	75.68	3400	$7.3 \cdot 10^5$	$7.3 \cdot 10^7$	$7.3 \cdot 10^9$

Table 2: λ_1, λ_2 are the eigenvalues of $\tilde{f}''(\alpha, \bar{\nu}, \bar{\kappa})$

We see that the quotient of the two eigenvalues explodes if the constant α increases. It is obvious that a minimizer $(\bar{\nu}, \bar{\kappa})$ for (10) is given by

$$\bar{\nu} = \bar{\mu} = 3.91, \ \bar{\kappa} = \frac{\overline{K}}{\alpha} = \alpha^{-1} 1.788E - 2. \tag{12}$$

Calculations show that any step length $h \leqslant 0.2$ is appropriate for $\alpha = 0.01$ in the gradient method applied to the least square process for (8). This is to be compared with $h = 10^{-5}$ for $\alpha = 1$. Column 3 in Table 2 shows that the problem is stiff for $\alpha = 1$ and unstiff if $\alpha = 0.01$.

2.4. The reader might think that the conjugate gradient method in its preconditioned form (compare [9])

$$x(t+h) - x(t) = \gamma(t) p(x(t))$$

$$p(x(t)) = -W^{-1}(x(t)) f'(x(t)) + \beta(t-h) p(x(t-h)) \tag{13}$$

$$\gamma(t) = \frac{f'(x(t))^T W^{-1}(x(t)) f'(x(t))}{p(x(t))^T f''(x(t)) p(x(t))}$$

$$\beta(t-h) = \frac{f'(x(t))^T W^{-1}(x(t)) f'(x(t))}{f'(x(t-h))^T W^{-1}(x(t-h)) f'(x(t-h))}$$

would be more appropriate to deal with the problem outlined in 2.3. This process may be put in our standard form

$$x(t+h) = x(t) + hV(h, x(t), x(t-h)), \tag{14}$$

$$V(h, x(t), x(t-h)) = -W^{-1}(x(t)) f'(x(t)) + x(t) - x(t-h)$$

if we assume $\gamma(t) = \beta(t) = h$ for all grid points t. Hence, if we simplify and put the functions $\gamma(t), \beta(t)$ constant, our Theorem 1.2 says that the process (13) has a continuous extension which is given by the differential equation

$$\dot{x}(t) = -W^{-1}(x(t)) f'(x(t)). \tag{15}$$

Actually, (14) is an explicit two-step method to solve the differential equation (15). An explicit procedure is never able to solve stiff equations in a satisfactory way. Therefore, we expect problems even for the preconditioned conjugate gradient method if stiffness is more and more pronounced. Table 3 gives our results:

Method	α_{max}
Conjugate gradients	0.03
Preconditioned conj. gradients	40
Explicit two-step-methods	≈ 2
A-stable ODE solver	> 1000

Table 3

The first column in this table identifies different methods to solve our least-square problem: A conjugate gradient method without preconditioning is (13) with $W(x(t)) \equiv I$. As an A-stable ODE solver (last row in Table 3), we have used the NAG-routine D02EBF. In the second column, we give the maximal constant α such that the corresponding least square problem could be successfully treated by the method in the first row. For example, the conjugate gradient method was successful for all constants $0 < \alpha \leqslant 0.03$. We see that even the preconditioned conjugate gradient method is unsuccessful if $\alpha > 40$, whereas the A-stable ODE solver D02EBF of the NAG library could manage values of α which exceed 1000. With more and more stiffness, however, the arithmetic of the computer comes into the game and plays an essential role for the success even of the A-stable ODE solver.

3. Boundary value problems

3.1. In [5] we have considered the system

$$p\Big(x_3(t) + \frac{h}{2}\Big)(x_1(t+h) - x_1(t)) = hx_2(t)$$

$$x_2(t+h) - x_2(t) = -hf(x_1(t+h)) \tag{1}$$

$$x_3(t+h) - x_3(t) = h$$

$$t = a + jh,\ j = 0, \ldots, \sigma(h) - 1$$

along with the sight conditions

$$\alpha_\tau x_1(\tau) - \beta_\tau x_2(\tau) = \gamma_\tau(x_1(a), x_2(a), x_1(b), x_2(b)),\ \tau = a, b \tag{2}$$

$$\alpha_a > 0,\ \alpha_b < 0,\ \beta_\tau \geqslant 0,\ \tau = a, b$$

and the initial condition

$$x_3(a) = a. \tag{3}$$

As in [5], we assume

$$f, p \in C(\mathbf{R}),\ p(t) > 0,\ 0 \leqslant f(\eta) \leqslant M,\ \eta \in \mathbf{R} \tag{4}$$

$$\gamma_\tau \in C(\mathbf{R}^4, \mathbf{R}),\ |\gamma_\tau(z)| \leqslant M,\ z \in \mathbf{R}^4,\ \tau = a, b$$

as well as

$$b = a + \sigma(h)h \tag{5}$$

for convenience. It is clear from [5] that (1) can be written as

$$p\Big(t + \frac{h}{2}\Big)(x_1(t+h) - x_1(t)) = hx_2(t)$$

$$x_2(t+h) - x_2(t) = -hf(x_1(t+h)) \tag{6}$$

$$t = a + jh,\ j = 0, \ldots, \sigma(h) - 1.$$

Furthermore, we have shown that the assumptions of Theorem 1.2 are satisfied for (1) if we complete the system with (2), (3). Moreover, the first component $v_1(t)$ of the function v guaranteed by Theorem 1.2 is a solution of the boundary value problem

$$(p(t)v_1'(t))' + f(v_1(t)) = 0 \ \text{ in } \ [a, b], \tag{7}$$

$$\alpha_\tau v_1(\tau) - \beta_\tau p(\tau) v_1'(\tau) = \gamma_\tau(v_1(a), p(a)v_1'(a), v_1(b), p(b)v_1'(b)),\ \tau = a, b.$$

3.2. In this Section we are going to consider a slightly perturbed system with about half as many equations as (6) and the same convergence properties as (1), (2), (3). This comes from the fact that any solution of (6) solves the equations

$$-p\Big(t - \frac{h}{2}\Big)x_1(t-h) + \Big(p\Big(t - \frac{h}{2}\Big) + p\Big(t + \frac{h}{2}\Big)\Big)x_1(t) - p\Big(t + \frac{h}{2}\Big)x_1(t+h) = \tag{8}$$

$$= h^2 f(x_1(t)),\ t = a + jh,\ j = 1, \ldots, \sigma(h) - 1$$

as we have shown in [5]. Conversely, any solution of (8) can be extended to a solution of (6) if we define $x_2(t)$ by the first equation in (6) for $t = a + jh$, $j = 0, \ldots, \sigma(h) - 1$ and by the second equation in (6) for $t = a + \sigma(h)h$. Therefore, it is enough to solve (8) along with (2). However, this system has two more unknowns $x_2(a)$, $x_2(b)$ than equations. A standard way out of this dilemma is a replacement of (2) by

$$\alpha_\tau x_1(\tau) - \beta_\tau z(\tau) = \gamma_\tau(x_1(a), z(a), x_1(b), z(b)), \tag{9}$$

$$2hz(\tau) := p(\tau)(x_1(\tau + h) - x_1(\tau - h)),\ \tau = a, b.$$

Now, x_2 is eliminated. However, the extra grid points $a - h, b + h$ appear. To find two extra equations we extend (8) and consider

$$(10)\quad -p\Big(t - \frac{h}{2}\Big)x_1(t - h) + \Big(p\Big(t - \frac{h}{2}\Big) + p\Big(t + \frac{h}{2}\Big)\Big)x_1(t) - p\Big(t + \frac{h}{2}\Big)x_1(t + h)$$

$$= h^2 f(x_1(t)),\ t = a + jh,\ j = 0, \ldots, \sigma(h).$$

The resulting system (9), (10) has as many scalar unknowns as it has equations. Obviously, this system has the same solutions as

$$p\Big(t + \frac{h}{2}\Big)(x_1(t + h) - x_1(t)) = hx_2(t)$$

$$(11)\quad x_2(t + h) - x_2(t) = -hf(x_1(t + h))$$

$$t = a + jh,\ j = -1, 0, 1, \ldots, \sigma(h)$$

along with (9). This system is of our standard form and Theorem 1.2 applies. To test the assumption (1.5), we note that we can write (9) as

$$\alpha_\tau x_1(\tau) - \beta_\tau \beta_\tau(h)x_2(\tau) = \gamma_\tau + \beta_\tau q_\tau(h, x_2(\tau), x_2(\tau - h))$$

$$(12)\quad \beta_\tau(h) = \frac{p(\tau)}{2}\left[\frac{1}{p(\tau + \frac{h}{2})} + \frac{1}{p(\tau - \frac{h}{2})}\right]$$

$$q_\tau(h, u, v) = \frac{p(\tau)}{2p(\tau - \frac{h}{2})}[v - u],\ \tau = a, a + \sigma(h)h.$$

This is a boundary condition very similar to (2) with (4) so that the reasoning in [5] to prove (1.5) for (1), (2), (3) takes over verbatim to (11), (12).

We are left with the transformation of (9) into (12): To this end we notice

$$2hz(\tau) = p(\tau)(x_1(\tau + h) - x_1(\tau) + x_1(\tau) - x_1(\tau - h))$$

$$(13)\quad = hp(\tau)\left[\frac{x_2(\tau)}{p(\tau + \frac{h}{2})} + \frac{x_2(\tau - h)}{p(\tau - \frac{h}{2})}\right]$$

$$= hp(\tau)\left[\frac{1}{p(\tau + \frac{h}{2})} + \frac{1}{p(\tau - \frac{h}{2})}\right]x_2(\tau) + h\frac{p(\tau)}{p(\tau - \frac{h}{2})}(x_2(\tau - h) - x_2(\tau))$$

$$= 2h\beta_\tau(h)x_2(\tau) + 2hq_\tau(h, x_2(t), x_2(t - \tau)).$$

Inserting the representation (13) into (9) yields (12).

The representation (12) also shows that, for $h \to 0$, the boundary conditions in (7) are obtained. We already proved in [5] that the differential equation in (7) results as the limit problem.

References

[1] *Boggs, P.T.*: The solution of nonlinear systems of equations by A-stable integration techniques, SIAM J. Numer. Anal. *8* (1971), 767–785.

[2] *Bohl, E.*: Finite Modelle gewöhnlicher Randwertaufgaben, Teubner Studienbücher, B.G. Teubner, 1981.

[3] *Bohl, E.*: Mathematische Grundlagen für die Modellierung biologischer Vorgänge, Springer Hochschultexte, Springer, 1987.

[4] *Bohl, E.*: Mathematik und Leben, Die Frage nach dem Leben, Serie Piper (Fischer, E.P., Mainzer, K., eds.), 1990, pp. 233–263.

[5] *Bohl, E.*: On the convergence of time-discrete processes, to appearin ZAMM 1993.

[6] *Collet, P., Eckmann, J.P.*: Iterated Maps on the Interval as dynamical Systems. Progress in Physics, Vol. 1, Basel.

[7] *Dahlquist, G.*: Convergence and stability in the numerical integration of ordinary differential equations, Math. Scand. *4* (1956), 33–53.

[8] *Dennis, J.E., Schnabel, R.B.*: Numerical Methods for Unconstrained Optimisation and Nonlinear Equations, Prentice-Hall Inc., Engelwood Cliffs, New Jersey, 1983.

[9] *Gill, P.E., Murray, W., Wright, M.H.*: Practical Optimization, Academic Press, London, New York, 1981.

[10] *Grigorieff, R.D.*: Numerik gewöhnlicher Differentialgleichungen 1, 2, Teubner Studienbücher, B.G. Teubner, 1972.

[11] *Hairer, E., Wanner, G., Norsett, P.S.*: Solving Ordinary Differential Equations I, Springer-Verlag, 1980.

[12] *May, R.*: Simple mathematical models with very complicated dynamics, Nature *261* (1976), 459–467.

[13] *Michaelis, L., Menten, M.L.*: Die Kinetik der Invertinwirkung, Biochem. Z. *49* (1913), 333–369.

[14] *Ortega, J.M., Rheinboldt, W.C.*: Iterative Solution of Nonlinear Equations in Several Variables, New York, San Francisco, London, 1970.

[15] *Schropp, J.*: Global dynamics of Newton's flow, in preparation.

[16] *Werner, J.*: Numerische Mathematik I, Vieweg, Braunschweig/Wiesbaden, 1992.

[17] *Wissel, C.*: Theoretische Ökologie, Springer, Berlin, Heidelberg, New York, 1989.

Authors' address: *Erich Bohl, Johannes Schropp*, Department of Mathematics, University of Konstanz, P.O. Box 5560, D-7750 Konstanz, Germany.