

Z. JACKIEWICZ (Fayetteville, Ark.) and M. KWAPISZ (Gdańsk)

THE NUMERICAL SOLUTION OF VOLTERRA INTEGRO-FUNCTIONAL EQUATIONS

Abstract. Methods for the numerical solution of the equation

$$z(t) = f\left(t, \int_{t_0}^t K(t, s, z(s))ds, z(\cdot)\right), \quad t \in [t_0, t_0 + a],$$

are proposed. These methods are compared with the method of successive approximations in the case of the equation $z(t) = f(t, z(\alpha(t)))$, $t \in [0, 1]$, $0 \leq \alpha(t) \leq t$. Some numerical examples are given.

1. Introduction. Let $C(I, R^n)$ denote the space of continuous functions defined on an interval $I = [t_0, t_0 + a]$, $t_0 \in R$, $a > 0$, with the uniform norm denoted by $\|\cdot\|^I$. For any $z \in C(I, R^n)$ and any subinterval $J \subset I$ we define a seminorm $\|z\|^J = \sup \{\|z(s)\|: s \in J\}$, where $\|\cdot\|$ stands for any norm in R^n . Let $I_t = [t_0, t]$ for $t \in I$. Consider the Volterra integro-functional equation

$$(1) \quad z(t) = f\left(t, \int_{t_0}^t K(t, s, z(s))ds, z(\cdot)\right), \quad t \in I,$$

where the functions f and K satisfy the following assumptions:

(H₁) The function $f: I \times R^n \times C(I, R^n) \rightarrow R^n$ is continuous and there exist non-negative constants L and M , $M < 1$, such that the Lipschitz condition

$$\|f(t, y_1, z_1(\cdot)) - f(t, y_2, z_2(\cdot))\| \leq L\|y_1 - y_2\| + M\|z_1 - z_2\|^I_t$$

holds for any $t \in I$, $y_1, y_2 \in R^n$, $z_1, z_2 \in C(I, R^n)$.

(H₂) The function $K: \Delta \times R^n \rightarrow R^n$, where $\Delta = \{(t, s): t \in I, t_0 \leq s \leq t\}$ is continuous and there exists a non-negative constant l such that

$$\|K(t, s, x_1) - K(t, s, x_2)\| \leq l\|x_1 - x_2\|$$

for any $(t, s) \in \Delta$ and $x_1, x_2 \in R^n$.

It is easy to prove that under the assumptions (H₁) and (H₂) the operator defined by the right-hand side of equation (1) is a contraction with respect to the norm

$$\|z\|^I = \sup \{\|z(s)\| \exp(-\lambda(s - t_0)): s \in I\}$$

(equivalent to the uniform norm) for any $\lambda > 0$ satisfying the inequality $M + L/\lambda < 1$. Clearly, this implies that equation (1) has a unique solution $Z \in C(I, R^n)$. Note also that if equation (1) is considered for $t \in [t_0, \infty)$, we can prove the existence and uniqueness result by the use of the same method.

Special cases of (1) are, among others, the Volterra integral equations of the second kind

$$(2) \quad z(t) = g(t) + \int_{t_0}^t K(t, s, z(s)) ds, \quad t \in I,$$

or

$$(3) \quad z(t) = f\left(t, \int_{t_0}^t K(t, s, z(s)) ds\right), \quad t \in I,$$

and functional equations in a much narrower sense considered by Kuczma [14].

Consider the Cauchy problem for the system of first-order differential equations

$$(4) \quad y'(t) = f(t, y(t)), \quad t \in I, \quad y(t_0) = y_0,$$

or, more generally,

$$(5) \quad y'(t) = f(t, y(t), y'(t)), \quad t \in I, \quad y(t_0) = y_0.$$

The problem (4) can be written in the form

$$(6) \quad y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds, \quad t \in I,$$

or

$$(7) \quad z(t) = f\left(t, y_0 + \int_{t_0}^t z(s) ds\right), \quad t \in I,$$

where

$$z(t) = y'(t), \quad y(t) = y_0 + \int_{t_0}^t z(s) ds.$$

Equations (6) and (7) are special cases of (2) and (3), respectively. It is also obvious that the problem (5) can be reduced to a special case of (1).

In the present paper we discuss the numerical solution of (1). Clearly, there exists an extensive literature concerned with the numerical solution of (4) (see, for instance, [9], [10], [15], and the references contained there) and there exists also a number of papers concerned with the numerical solution of integral equations (see [1]–[4], [6]–[8], [16]). But there exist only a few papers which deal with the numerical solution of the initial-value problems

for the so-called differential-delay equations of neutral type (see [5], [11]–[13], [17], [19]). Obviously, some classes of such equations can be easily reduced to special cases of (1). Finally, the functional equations of the form $z(t) = f(t, z(\cdot))$, $t \in I$, were not numerically solved to the present time.

These are the reasons why we try to discuss the numerical solution of (1) and to unify some methods used for the numerical treatment of special cases of this equation.

Let $h > 0$ denote the constant step size and put $t_i = t_0 + ih$, $i = 0(1)N$, where $Nh = a$. In order to solve the problem (1) numerically consider the following class of approximate methods:

$$(8) \quad z_h(t_{i+k-1} + rh) = \sum_{j=0}^k \beta_j(r) z_h(t_{i+j}), \quad r \in (0, 1),$$

$$(9) \quad z_h(t_{i+k}) = f_h(t_{i+k}, h \sum_{j=0}^{i+k} w_{i+k,j} K(t_{i+k}, t_j, z_h(t_j)), z_h(\cdot)),$$

$i = 0(1)N - k$. It is assumed that the approximate solution $z_h \in \tilde{C}^0(I, R^n)$ is given for $t \in [t_0, t_{k-1}]$. Here $\tilde{C}^0(I, R^n)$ stands for the space of piecewise continuous functions with discontinuities of the first kind. For $z \in \tilde{C}^0(I, R^n)$ and any subinterval J of I , $\|z\|^J$ is defined as previously. We assume that:

(H₃) The functions $\beta_j: [0, 1] \rightarrow R$, $j = 0(1)k$, are continuous and $\|\beta_k\|^{[0,1]} = 1$.

(H₄) The weights $w_{i,j}$, $i = 0(1)N$, $j = 0(1)i$, are uniformly bounded. This means that there exists a constant W , $0 \leq W < \infty$, such that $|w_{i,j}| \leq W$ for $i = 0(1)N$, $j = 0(1)i$.

(H₅) The operators $f_h: I \times R^n \times \tilde{C}^0(I, R^n) \rightarrow R^n$, $h \in I_{h_0} := [0, h_0]$, $h_0 > 0$, satisfy the Lipschitz condition (uniformly in h)

$$\|f_h(t, y_1, z_1(\cdot)) - f_h(t, y_2, z_2(\cdot))\| \leq L\|y_1 - y_2\| + M\|z_1 - z_2\|^{I_t}$$

for $t \in I$, $y_1, y_2 \in R^n$, $z_1, z_2 \in \tilde{C}^0(I, R^n)$. We may assume without loss of generality that the constants L and M are the same as in (H₁).

(H₆) The operators f_h , $h \in I_{h_0}$, are discrete. This means that f_h depends only on $z_h(t)$ for t belonging to some finite set T_h .

To explain (H₆) in slightly more detail, note that in the case of equations (2), (3), (6), and (7) we can simply put $f_h = f$. But we cannot do this if, for example,

$$\begin{aligned} f\left(t, \int_{t_0}^t K(t, s, z(s)) ds, z(\cdot)\right) \\ = f\left(t, \int_{t_0}^t K(t, s, z(s)) ds, z(t), z(t_0 + (t - t_0)/2), z(t_0 + (t - t_0)/3), \dots\right). \end{aligned}$$

Note that if the right-hand side of (9) depends only on t_j and $z_h(t_j)$ for $j = 0(1)i+k$, then formula (8) is superfluous when we are interested in approximation of the solution Z of (1) only on the grid $\{t_i\}_{i=0}^N$. This takes place in the case of the integral equation (2). Now we arrive at the method of the form

$$z_{i+k} = g(t_{i+k}) + h \sum_{j=0}^{i+k} w_{i+k,j} K(t_{i+k}, t_j, z_j),$$

$i = 0(1)N-k$, where $z_i = z_h(t_i)$. This class of methods was considered, for example, by Baker and Keech [3]. But this is not the case if equation (1) takes the form

$$(10) \quad z(t) = f(t, z(\alpha(t))), \quad t \in I,$$

where, for example, $\alpha(t) = t/2$ and $f_h(t, z_h(\cdot)) = f(t, z_h(t/2))$. The solving of the equation of type (10) is discussed in more detail in Section 5.

2. The convergence of the method. First of all we check that the method (8)–(9) determines the approximate solution z_h uniquely. To see this assume that z_h is given for $t \in [t_0, t_{i+k-1}]$ and consider the equations

$$(11) \quad q = F(q),$$

where

$$F(q) := f_h(t_{i+k}, h \sum_{j=0}^{i+k-1} w_{i+k,j} z_h(t_j) + h w_{i+k,i+k} q, z_q(\cdot))$$

and

$$(12) \quad z_q(t) = \begin{cases} z_h(t), & t \in [t_0, t_{i+k-1}], \\ \sum_{j=0}^{k-1} \beta_j(r) z_h(t_{i+j}) + q \beta_k(r), & t = t_{i+k-1} + rh, r \in (0, 1), \\ q, & t = t_{i+k}. \end{cases}$$

In view of (H₂), (H₃), and (H₅), we have

$$\|F(q_1) - F(q_2)\| \leq (hWL + M \|\beta_k\|^{[0,1]}) \|q_1 - q_2\| = (hWL + M) \|q_1 - q_2\|.$$

Thus F is a contraction for h sufficiently small to imply that equation (11) has a unique solution q . Consequently, by the use of (12), z_h can be defined on the interval $[t_0, t_{i+k}]$ and, by induction, on the whole interval I .

Define the local error function of the method (8)–(9) at the point $t + (k-1+r)h$ by

$$v(t, r, h) := Z(t + (k-1+r)h) - \sum_{j=0}^k \beta_j(r) Z(t + jh).$$

Define also

$$\xi(t, h) := f\left(t, \int_{t_0}^t K(t, s, Z(s))ds, Z(\cdot)\right) - f_h\left(t, \int_{t_0}^t K(t, s, Z(s))ds, Z(\cdot)\right),$$

$$\eta(t, h) := \int_{t_0}^t K(t, s, Z(s))ds - h \sum_{j=0}^i w_{i,j} K(t_i, t_j, Z(t_j)),$$

where i is the largest integer such that $ih \leq t - t_0$, and put

$$v(h) := \sup \{ \|v(t, r, h)\| : t \in [t_0, t_0 + a - kh], r \in [0, 1] \},$$

$$\xi(h) := \sup \{ \|\xi(t, h)\| : t \in I \}, \quad \eta(h) := \sup \{ \|\eta(t, h)\| : t \in I \}.$$

The method (8)–(9) is said to be *consistent* (with (1) on the solution Z) if $v(h) = o(1)$, $\xi(h) = o(1)$, and $\eta(h) = o(1)$ as $h \rightarrow 0$. The *order of consistency* is p (or, simply, the *method is of order p*) if $v(h) = O(h^p)$, $\xi(h) = O(h^p)$, and $\eta(h) = O(h^p)$ as $h \rightarrow 0$.

Define the global error function of the method (8)–(9) by $e_h = Z - z_h$.

The method (8)–(9) is said to be *convergent* (to the solution Z of (1)) if $\|e_h\|^l \rightarrow 0$ as $h \rightarrow 0$. The *order of convergence* is p if $\|e_h\|^l = O(h^p)$ as $h \rightarrow 0$.

Now we can prove the main result:

THEOREM. Assume that

1° the conditions (H_1) – (H_5) hold;

2° the method (8)–(9) is consistent;

3° $DM < 1$, where $D = \sup \left\{ \sum_{j=0}^k |\beta_j(r)| : r \in [0, 1] \right\}$;

4° $\|e_h\|^{[l_0 \cdot k - 1]} = o(1)$ as $h \rightarrow 0$.

Then the method (8)–(9) is convergent to the solution Z of (1).

Proof. Subtracting (8) from the equation

$$Z(t_{i+k-1} + rh) = \sum_{j=0}^k \beta_j(r) Z(t_{i+j}) + v(t_i, r, h),$$

we obtain

$$e_h(t_{i+k-1} + rh) = \sum_{j=0}^k \beta_j(r) e_h(t_{i+j}) + v(t_i, r, h),$$

$i = 0(1)N - k$, $r \in (0, 1)$. Taking norms in this equation, in view of 3°, we get

$$(13) \quad \|e_h\|^{[l_0 \cdot k - 1, l_0 \cdot k]} \leq D \max \{ \|e_h(t_{i+j})\| : j = 0(1)k \} + v(h),$$

$i = 0(1)N - k$. Subtracting the equations

$$Z(t_{i+k}) = \left(t_{i+k}, \int_{t_0}^{t_{i+k}} K(t_{i+k}, s, Z(s))ds, Z(\cdot) \right),$$

$$z_h(t_{i+k}) = f_h\left(t_{i+k}, h \sum_{j=0}^{i+k} w_{i+k,j} K(t_{i+k}, t_j, z_h(t_j)), z_h(\cdot)\right),$$

$i = 0(1)N - k$, and using (H_3) we obtain

$$\|e_h(t_{i+k})\| \leq L \left\| \int_{t_0}^{t_{i+k}} K(t_{i+k}, s, Z(s)) ds - h \sum_{j=0}^{i+k} w_{i+k,j} K(t_{i+k}, t_j, z_h(t_j)) \right\| + \\ + M \|e_h\|^{[t_0, t_{i+k}]} + \zeta(h),$$

$i = 0(1)N - k$. Hence

$$(14) \quad \|e_h(t_i)\| \leq \|e_h\|^{[t_0, t_k-1]} + \\ + L \left\| \int_{t_0}^{t_i} K(t_i, s, Z(s)) ds - h \sum_{j=0}^i w_{i,j} K(t_i, t_j, z_h(t_j)) \right\| + \\ + M \|e_h\|^{[t_0, t_i]} + \zeta(h)$$

for all $i = 0(1)N$. Now we estimate the second term on the right-hand side of inequality (14). We have

$$\left\| \int_{t_0}^{t_i} K(t_i, s, Z(s)) ds - h \sum_{j=0}^i w_{i,j} K(t_i, t_j, z_h(t_j)) \right\| \\ \leq \left\| \int_0^{t_i} K(t_i, s, Z(s)) ds - h \sum_{j=0}^i w_{i,j} K(t_i, t_j, Z(t_j)) \right\| + \\ + \left\| h \sum_{j=0}^i w_{i,j} K(t_i, t_j, Z(t_j)) - h \sum_{j=1}^i w_{i,j} K(t_i, t_j, z_h(t_j)) \right\| \\ \leq \eta(h) + hWl \sum_{j=0}^i \|e_h(t_j)\| \leq \eta(h) + hWl \sum_{j=0}^i \|e_h\|^{[t_0, t_j]}$$

and (14) takes the form

$$\|e_h(t_i)\| \leq \|e_h\|^{[t_0, t_k-1]} + hWl \sum_{j=0}^i \|e_h\|^{[t_0, t_j]} + M \|e_h\|^{[t_0, t_i]} + L\eta(h) + \zeta(h),$$

$i = 0(1)N$. Substitution of this inequality into (13) yields

$$\|e_h\|^{[t_{i+k-1}, t_{i+k}]} \leq D (\|e_h\|^{[t_0, t_k-1]} + hWl \sum_{j=0}^{i+k} \|e_h\|^{[t_0, t_j]} + \\ + M \|e_h\|^{[t_0, t_{i+k}]} + L\eta(h) + \zeta(h)) + v(h),$$

$i = 0(1)N - k$. But the sequence defined by the right-hand side of the last inequality is non-decreasing in i . Hence

$$\|e_h\|^{[t_0, t_{i+k}]} \leq D^* \|e_h\|^{[t_0, t_k-1]} + hDWl \sum_{j=0}^{i+k} \|e_h\|^{[t_0, t_j]} + \\ + DM \|e_h\|^{[t_0, t_{i+k}]} + DL\eta(h) + D\zeta(h) + v(h),$$

$i = 0(1)N - k$, where $D^* = \max\{1, D\}$. Let $h^* \in I_{h_0} - \{0\}$ be such that $1 - DM - h^*DWL > 0$. Putting

$$A = (1 - DM - h^*DWL)^{-1},$$

$$B = AD^*,$$

$$C = ADWL,$$

$$\omega(h) = ADL\eta(h) + AD\xi(h) + Av(h),$$

after simple calculations we obtain

$$\|e_h\|^{[t_0, t_{i+k}]} \leq B \|e_h\|^{[t_0, t_{k-1}]} + hC \sum_{j=0}^{i+k-1} \|e_h\|^{[t_0, t_j]} + \omega(h),$$

$i = 0(1)N - k$. Denote the right-hand side of this inequality by u_i . Then $\|e_h\|^{[t_0, t_{i+k}]} \leq u_i$, $i = 0(1)N - k$, and $u_{i+1} \leq (1 + hC)u_i$, $i = 0(1)N - k - 1$. By induction it follows that $u_i \leq (1 + hC)^i u_0 \leq \exp(Ca)u_0$, $i = 0(1)N - k$. Finally, we obtain the inequality

$$(15) \quad \|e_h\|^I \leq (B^* \|e_h\|^{[t_0, t_{k-1}]} + \omega(h)) \exp(Ca),$$

where $B^* = B + Ckh_0$, and the theorem follows.

Remark. It is easy to see that if $\|e_h\|^{[t_0, t_{k-1}]} = O(h^p)$ as $h \rightarrow 0$ and if the method (8)–(9) is of order p , then the order of convergence is also p .

3. Construction of methods. Tavernini [18] has demonstrated a way of constructing linear multistep methods of any order for the numerical solution of functional differential equations. We adopt his approach here. Assume that $Z \in C^{k+2}(I, R^n)$, where $C^i(I, R^n)$ is the space of functions with continuous i -th derivative. After simple calculations we obtain

$$\begin{aligned} v(t, r, h) = & \sum_{l=0}^k (h^l/l!) [(k+r-1)^l - \sum_{j=0}^k \beta_j(r)j^l] Z^{(l)}(t) + \\ & + (h^{k+1}/(k+1)!)[(k+r-1)^{k+1} - \sum_{j=0}^k \beta_j(r)j^{k+1}] Z^{(k+1)}(t) + \\ & + O(h^{k+2}). \end{aligned}$$

If we want to obtain a method of order $k+1$, we must require $v(h) = O(h^{k+1})$, $\xi(h) = O(h^{k+1})$, and $\eta(h) = O(h^{k+1})$. The first requirement leads to the system of equations

$$\sum_{j=0}^k \beta_j(r)j^l = (k+r-1)^l,$$

$l = 0(1)k$, which uniquely determines the functions β_j , $j = 0(1)k$, because the determinant of this system is the Vandermonde determinant. We also have the estimate

(16)

$$v(h) \leq (h^{k+1}/(k+1)!) \sup \left\{ \left| (k+r-1)^{k+1} - \sum_{j=0}^k \beta_j(r) j^{k+1} \right| : r \in [0, 1] \right\} Z_*^{(k+1)} + O(h^{k+2}),$$

where

$$Z_*^{(k+1)} := \sup \{ \|Z^{(k+1)}(t)\| : t \in I \}.$$

The second requirement, $\xi(h) = O(h^{k+1})$, is fulfilled if we put, for example, $f_h = f$. But it is not always possible to do so because the operators f_h must also be discrete. This point was mentioned earlier in Section 1. The third requirement, $\eta(h) = O(h^{k+1})$, is satisfied if we choose a quadrature formula of order $k+1$. For $k=1$ we can take the repeated trapezium rule. For $k=2$ we may combine the repeated Simpson's rule with the trapezium rule (as a result we obtain a quadrature formula of order 3). For $k=3$ we may combine the repeated Simpson's rule with the (3/8)'ths rule (as a result we obtain a quadrature formula of order 4). All these quadrature rules are given, for example, in [1] and [3].

The resulting methods, for $k=1, 2$, and 3 , and the estimates of $v(h)$ obtained from (16) are listed below:

$k=1$:

$$(17) \quad \begin{aligned} z_h(t_i + rh) &= (1-r)z_h(t_i) + rz_h(t_{i+1}), \quad r \in (0, 1), \\ z_h(t_{i+1}) &= f_h(t_{i+1}, h \sum_{j=0}^{i+1} w_{i+1,j} K(t_{i+1}, t_j, z_h(t_j)), z_h(\cdot)), \end{aligned}$$

$i = 0(1)N-1$, where

$$w_{s,j} = \begin{cases} 1/2, & j = 0, s, \\ 1, & j = 1(1)s-1. \end{cases}$$

Now (16) takes the form

$$v(h) \leq (h^2/8) Z_*'' + O(h^3), \quad h \rightarrow 0.$$

$k=2$:

$$(18) \quad \begin{aligned} z_h(t_{i+1} + rh) &= (r^2/2 - r/2)z_h(t_i) + (1-r^2)z_h(t_{i+1}) + \\ &\quad + (r/2 + r^2/2)z_h(t_{i+2}), \quad r \in (0, 1), \\ z_h(t_{i+2}) &= f_h(t_{i+2}, h \sum_{j=0}^{i+2} w_{i+2,j} K(t_{i+2}, t_j, z_h(t_j)), z_h(\cdot)), \end{aligned}$$

$i = 0(1)N-2$, where

$$(19) \quad w_{2s,j} = \begin{cases} 1/3, & j = 0, 2s, \\ 4/3, & j = 1(2)2s-1, \\ 2/3, & j = 2(2)2s-2, \end{cases}$$

$$w_{2s+1,j} = \begin{cases} 1/2, & j = 0, \\ 1/3 + 1/2, & j = 1, \\ 4/3, & j = 2(2)2s, \\ 1/3, & j = 2s+1, \end{cases}$$

or, preferably, (19) together with

$$w_{2s+1,j} = \begin{cases} 1/3, & j = 0, \\ 4/3, & j = 1(2)2s-1, \\ 1/3 + 1/2, & j = 2s, \\ 1/2, & j = 2s+1 \end{cases}$$

(cf. [1] and [3]). The estimate (16) takes now the form

$$v(h) \leq (3^{1/2} h^3/27) Z_*^{(3)} + O(h^4), \quad h \rightarrow 0.$$

$k = 3$:

$$(20) \quad \begin{aligned} z_h(t_{i+2} + rh) = & (r/6 - r^3/6) z_h(t_i) + (-r + r^2/2 + r^3/2) z_h(t_{i+1}) + \\ & + (1 + r/2 - r^2 - r^3/2) z_h(t_{i+2}) + (r/2 + r^2/2 + \\ & + r^3/6) z_h(t_{i+3}), \quad r \in (0, 1), \end{aligned}$$

$$z_h(t_{i+3}) = f_h(t_{i+3}, h \sum_{j=0}^{i+3} w_{i+3,j} K(t_{i+3}, t_j, z_h(t_j)), z_h(\cdot)),$$

$i = 0(1)N-3$, where $w_{2s,j}$ is given by (19) and

$$w_{2s+1,j} = \begin{cases} 3/8, & j = 0, \\ 9/8, & j = 1, 2, \\ 3/8 + 1/3, & j = 3, \\ 4/3, & j = 4(2)2s, \\ 2/3, & j = 5(2)2s-1, \\ 1/3, & j = 2s+1, \end{cases}$$

or, preferably, (19) together with

$$w_{2s+1,j} = \begin{cases} 1/3, & j = 0, \\ 4/3, & j = 1(2)2s-3, \\ 2/3, & j = 2(2)2s-4, \\ 1/3 + 3/8, & j = 2s-2, \\ 9/8, & j = 2s-1, 2s, \\ 3/8, & j = 2s+1. \end{cases}$$

In this case (16) takes the form

$$v(h) \leq (2h^4/3) Z_*^{(4)} + O(h^5), \quad h \rightarrow 0.$$

Remark. The methods listed above correspond to the methods of Adams–Moulton type for neutral functional differential equations considered in [11], where $b'_j = \beta_j$, $j = 0(1)k$.

4. The case of ordinary differential equations. Consider the problem

$$(21) \quad y'(t) = f(t, y(t)), \quad t \in I, \quad y(t_0) = 0.$$

Putting

$$y(t) = \int_{t_0}^t z(s) ds,$$

we can write this equation in the form

$$(22) \quad z(t) = f\left(t, \int_{t_0}^t z(s) ds\right), \quad t \in I.$$

On the basis of (8)–(9) we can define a method for the numerical solution of the problem (22). This method takes the form

$$(23) \quad z_{i+k} = f\left(t_{i+k}, h \sum_{j=0}^{i+k} w_{i+k,j} z_j\right),$$

$i = 0(1)N-k$, where z_j , $j = 0(1)k-1$, are given. As a result of applying (23) to (22), we obtain an approximation to Z , the solution of (22), on the grid $\{t_i\}_{i=0}^N$. But when solving the problem (21), we are interested in an approximation to Y , the solution of (21), rather than to Z . But the convergent method for the numerical solution of the problem (21) can be defined as follows:

$$(24) \quad z_{i+k} = f\left(t_{i+k}, h \sum_{j=0}^{i+k} w_{i+k,j} z_j\right), \quad y_{i+k} = h \sum_{j=0}^{i+k} w_{i+k,j} z_j,$$

$i = 0(1)N-k$, where y_j , $j = 0(1)k-1$, are given and $z_j = f(t_j, \tilde{y}_j)$, $j = 0(1)k-1$. It is obvious that if $y_j \rightarrow Y(t_0)$, $j = 0(1)k-1$, then $z_j \rightarrow Z(t_0) = f(t_0, Y(t_0))$ with the same order. It follows from the previous considerations that

$$\sup \{\|e_j\| : j = 0(1)N\} \rightarrow 0 \quad \text{as } h \rightarrow 0,$$

where $e_j := z_j - Z(t_j)$. Define $\varepsilon_i := y_i - Y(t_i)$, $i = 0(1)N$. Then

$$\begin{aligned}
\|\varepsilon_{i+k}\| &= \left\| h \sum_{j=0}^{i+k} w_{i+k,j} z_j - \int_{t_0}^{t_{i+k}} Z(s) ds \right\| \\
&\leq \left\| h \sum_{j=0}^{i+k} w_{i+k,j} z_j - h \sum_{j=0}^{i+k} w_{i+k,j} Z(t_j) \right\| + \left\| h \sum_{j=0}^{i+k} w_{i+k,j} Z(t_j) - \int_{t_0}^{t_{i+k}} Z(s) ds \right\| \\
&\leq hW \sum_{j=0}^{i+k} \|e_j\| + \eta(h) \leq Wa \sup \{\|e_j\|: j = 0(1)N\} + \eta(h),
\end{aligned}$$

$i = 0(1)N - k$. Finally,

$$\begin{aligned}
&\sup \{\|\varepsilon_j\|: j = 0(1)N\} \\
&\leq \sup \{\|\varepsilon_j\|: j = 0(1)k-1\} + Wa \sup \{\|e_j\|: j = 0(1)N\} + \eta(h),
\end{aligned}$$

and the convergence of the method (24) to the solution Y of (21) follows. It is also obvious that if the method (23) is of order p , then the method (24) is also of order p .

It is also well known that the methods of type (8)–(9) applied to the integral equation

$$y(t) = \int_{t_0}^t f(s, y(s)) ds$$

being equivalent to the problem (21) converge (cf. [2]). However, in the case of the Cauchy problem (21) we have a number of numerical methods which are more effective than those discussed in this section.

5. Other special cases of equation (1). Consider now a special case of (1), namely

$$(25) \quad z(t) = f(t, z(\alpha(t))), \quad t \in [0, 1],$$

where the function $\alpha: [0, 1] \rightarrow [0, 1]$ is continuous and $0 \leq \alpha(t) \leq t$ for $t \in [0, 1]$. The condition (H_1) now takes the form

$$\|f(t, z_1) - f(t, z_2)\| \leq L \|z_1 - z_2\|$$

with $0 \leq L < 1$, for $t \in [0, 1]$, $z_1, z_2 \in R^n$.

We now explain how the method (8)–(9) works in this case. Let us assume that the approximate solution z_h is given on the interval $[t_0, t_{k-1}]$ and on the grid $\{t_{s+k}\}_{s=0}^{i-1}$ and try to compute $z_h(t_{i+k})$. In view of (9), for $f_h = f$ we have

$$(26) \quad z_h(t_{i+k}) = f(t_{i+k}, z_h(\alpha(t_{i+k}))),$$

which together with (8) comprises the method for equation (25). The following three cases are possible:

1. $\alpha(t_{i+k}) \in [t_0, t_{k-1}]$. Then $z_h(t_{i+k})$ can be computed directly from equation (26).

2. $\alpha(t_{i+k}) \in (t_{s+k-1}, t_{s+k})$ for some integer s , $0 \leq s < i$. Then, in view of (8),

$$z_h(\alpha(t_{i+k})) = \sum_{j=0}^k \beta_j(r(s)) z_h(t_{s+j}),$$

where $r(s) \in (0, 1)$ is defined by $t_{s+k-1} + r(s)h = \alpha(t_{i+k})$. Now $z_h(t_{i+k})$ can be computed from the equation

$$z_h(t_{i+k}) = f(t_{i+k}, \sum_{j=0}^k \beta_j(r(s)) z_h(t_{s+j})).$$

Note that if $s = i$, this equation is implicit.

3. $\alpha(t_{i+k}) = t_{s+k}$ for some integer s , $0 \leq s \leq i$. Then $z_h(t_{i+k})$ can be computed from the equation

$$z_h(t_{i+k}) = f(t_{i+k}, z_h(t_{s+k})).$$

Summing up the above discussion, note that the method (8)–(9) applied to equation (25) can be written in the form

$$(27) \quad z_h(t_{i+k}) = \begin{cases} f(t_{i+k}, z_h(\alpha(t_{i+k}))), & \alpha(t_{i+k}) \in [t_0, t_{k-1}], \\ f(t_{i+k}, \sum_{j=0}^{i+k} c_{i+k,j} z_h(t_j)), & \alpha(t_{i+k}) \notin [t_0, t_{k-1}], \end{cases}$$

$$(28) \quad z_h(t_{i+k-1} + rh) = \sum_{j=0}^k \beta_j(r) z_h(t_{i+j}), \quad r \in (0, 1),$$

$i = 0(1)N-k$, where $z_h|_{[t_0, t_{k-1}]}$ is given and the coefficients $c_{i+k,j}$ depend on α_j , β_j and t_{i+k} . Note in passing that at most $k+1$ coefficients $c_{i+k,j}$ differ from 0 and that $|c_{i+k,i+k}| \leq 1$. The last condition follows from (H_3) and the equality $c_{i+k,i+k} = \beta_k(r(i))$, and it ensures that equation (27) has a unique solution $z_h(t_{i+k})$ (see also the beginning of Section 2).

Since all assumptions of the convergence theorem from Section 2 hold, assuming that the starting error $\|e_h\|^{[t_0, t_{k-1}]}$ is negligible and putting $W = 0$, we obtain, in view of (15), the following estimate for the global error function of the method (27)–(28):

$$\|e_h\|^{[0,1]} \leq v(h)/(1-DL).$$

For the methods considered in Section 3 we get

$$\|e_h\|^{[0,1]} \leq B_k h^{k+1} + O(h^{k+2}), \quad h \rightarrow 0,$$

where, for example,

$$B_1 = Z''_*/(8(1-DL)), \quad B_2 = Z^{(3)}_*/(9(1-DL)), \quad B_3 = 2Z^{(4)}_*/(3(1-DL)).$$

On the other hand, equation (25) can also be solved by the method of successive approximations. This method takes the form

$$(29) \quad z_0(t) = f(t, z_*), \quad z_{n+1}(t) = f(t, z_n(\alpha(t))),$$

$n = 0, 1, \dots$, where z_* is arbitrary and fixed. We can take for z_* the unique solution of the equation $z = f(0, z)$. It can be proved by induction that

$$\|Z(t) - z_n(t)\| \leq L^{n+1} \|Z(\alpha^{n+1}(t)) - z_*\|$$

for $t \in [0, 1]$. A rougher estimate can be given by

$$\|Z(t) - z_n(t)\| \leq KL^{n+1},$$

where $K = \|z_*\| + \sup \{\|Z(t)\| : t \in [0, 1]\}$. It follows from the last inequality, in view of $L < 1$, that the sequence z_n ($n = 0, 1, \dots$) converges to the solution Z of (25) uniformly in t .

Suppose now that we want to compute an approximation to $Z(1)$ using the methods (27)–(28) and (29). We attempt to compare the efficiency of these methods assuming that

$$(30) \quad KL^{n+1} = B_k h^{k+1},$$

i.e. that the estimates of the global error are equal for both methods. By *efficiency* we mean the number of evaluations of the function f . We denote this number by E_1 and E_2 for the methods (27)–(28) and (29), respectively. If the method (27)–(28) is used, we must usually solve at each step of the calculations a nonlinear equation given by (27). Assume that this equation is solved by the method of successive approximations. In order to estimate the required number of iterations at each step, rewrite (27) in the form $x = F(x)$ with an appropriate function F . In order not to destroy the accuracy of the method, the iterations must be carried out until $\|x_m - x_*\| \leq \beta h^{k+1}$, where x_* is the fixed point of the function F and β is some constant. In practice, the iterations are carried out until two successive iterations differ by less than βh^{k+1} . We have $\|x_m - x_{m-1}\| \leq L^{m-1} \|F(x_0) - x_0\|$ and we require that $L^{m-1} X \leq \beta h^{k+1}$, where $X = \|F(x_0) - x_0\|$. Hence the number of iterations m satisfies the inequality

$$m \geq \varepsilon \log_L(\beta h^{k+1}/X) + 1,$$

where $\varepsilon = 0$ if $\log_L(\beta h^{k+1}/X) \leq 0$, and $\varepsilon = 1$ if $\log_L(\beta h^{k+1}/X) > 0$. Assuming that in order to get the starting values $z_h(t_s)$, $s = 0(1)k-1$, we also need m iterations at each step, we obtain

$$E_1 = (N+1)m = [\varepsilon \log_L(\beta h^{k+1}/X) + 1](1 + 1/h).$$

In the case of the method (29), if we accept $z_n(1)$ as a final approximation to $Z(1)$ and if we assume that m iterations are needed in order to get z_* , we

obtain

$$E_2 = m + n + 1 = \varepsilon \log_L(\beta h^{k+1}/X) + n + 2.$$

Now we compare E_1 and E_2 under the assumption (30). Assuming $K > 0$ and putting

$$C = (B_k/K)^{1/(k+1)}, \quad Q_\varepsilon = \varepsilon \log_L(\beta K/B_k X) + \varepsilon + 1, \quad \varepsilon = 0, 1,$$

after simple calculations we obtain

$$E_1 = CL^{-(n+1)/(k+1)}(\varepsilon n + Q_\varepsilon) + \varepsilon n + Q_\varepsilon \quad \text{and} \quad E_2 = (1 + \varepsilon)n + Q_\varepsilon + 1.$$

The formulas for E_1 and E_2 were derived for two extreme cases, where $\varepsilon = 1$ or $\varepsilon = 0$ for all grid points. In practice rather intermediate cases take place. This means that for some grid points $\varepsilon = 0$, and for others $\varepsilon = 1$. In any case, it is easy to see that $E_1 > E_2$ for n sufficiently large.

Assume now that we want to compute an approximation to Z on the whole grid $\{t_i\}_{i=0}^N$. Then E_1 does not change and E_2 is given by

$$E_2 = m + N(n+1) = CL^{-(n+1)/(k+1)}(n+1) + \varepsilon n + Q_\varepsilon.$$

In this case we cannot compare the efficiency of the considered methods without examining the size of the constant Q_ε depending on equation (25) and on the choice of the points x_0 . If these starting points were chosen properly so that $X = \|F(x_0) - x_0\|$ are sufficiently small, then $\log_L(\beta h^{k+1}/X) \leq 0$. This corresponds to the second extreme case, $\varepsilon = 0$. Now

$$E_1 = CL^{-(n+1)/(k+1)} + 1, \quad E_2 = CL^{-(n+1)/(k+1)}(n+1) + 1,$$

and $E_2 > E_1$ for all n .

The efficiency analysis given above is purely qualitative. In the next section we compare E_1 and E_2 for a concrete example, but we do this in a somewhat different way: we compare E_1 and E_2 when the real errors of the methods discussed are almost equal.

6. Numerical example. Consider the linear functional equation

$$z(t) = (1/3) \cos(t/2) z(t/2) + (5/6) \sin(t), \quad t \in [0, 1],$$

with the theoretical solution $z(t) = \sin(t)$.

This example was solved by the methods (17), (18), and (20) (we denote these methods by AM2, AM3, and AM4, respectively) and by the method (29), where $z_h(0) = z_* = z(0) = 0$. The computations were carried out on the Odra 1204 computer at the Institute of Mathematics, University of Gdańsk. The results of computations by the method AM2 are given in Table 1, where $e_h(1) := |Z(1) - z_h(1)|$, and by the method (29) in Table 2, where $e_i(1) := |T(1) - z_i(1)|$.

TABLE 1

h	$e_h(1) \cdot 10^9$	time (sec)	E_{teor}	E_{real}
2^{-1}	4 501 238	1.107	3	2
2^{-2}	647 613	1.161	6	4
2^{-3}	19 444	1.304	10	9
2^{-4}	3 065	1.520	19	17
2^{-5}	503	1.951	36	33
2^{-6}	14	2.849	68	66
2^{-7}	2	4.577	133	130
2^{-8}	0	8.067	262	259

TABLE 2

i	$e_i(1) \cdot 10^9$	time (sec)	$E_2 = n + 1$
1	23 374 194	0.056	2
2	3 895 699	0.085	3
3	649 283	0.115	4
4	108 214	0.143	5
5	18 036	0.173	6
6	3 006	0.202	7
7	501	0.232	8
8	83	0.262	9
9	14	0.291	10
10	2	0.321	11
11	0	0.350	12

The number E_{teor} in Table 1 is equal to $m + N - 1$, where m is the estimated number of iterations needed for computation of the value $z_n(t_1)$. This number was calculated in the previous section and is given by

$$m = \varepsilon \log_L(\beta h^2/X) + 1.$$

To evaluate X , note that $z_h(t_1)$ is computed from the equation

$$z_h(t_1) = (1/3) \cos(t_1/2) \cdot (1/2)(z_h(t_0) + z_h(t_1)) + (5/6) \sin(t_1),$$

where as the first approximation we take $z_h^{[0]}(t_1) = z_h(t_0) = 0$. Then it is easy to see that $X = (5/6) \sin(h)$. We may put $L = 1/3$ and assume that $\beta = 5/6$. Then

$$E_{\text{teor}} = \varepsilon \log_{1/3}(h^2/\sin(h)) + N.$$

The number E_{real} is equal to $m^* + N - 1$, where m^* is the smallest integer such that

$$|z_h^{[m^*]}(t_1) - z_h^{[m^*-1]}(t_1)| \leq (5/6) h^2.$$

Note that for $h = 2^{-6}$ and for $n = 9$ we received the same error for the methods (17) and (29). We see from Tables 1 and 2 that the iterative method (29) is more efficient than the method (17).

The results obtained for the methods AM3 and AM4 and the step sizes given in Table 1 coincide with the exact solution to the number of nine digits.

References

- [1] C. T. H. Baker, *The numerical treatment of integral equations*, Clarendon Press, Oxford 1977.
- [2] — *Initial value problems for Volterra integro-differential equations*, pp. 296–307 in: *Modern numerical methods for ordinary differential equations* (G. Hall and J. M. Watt, eds.), Clarendon Press, Oxford 1976.
- [3] — and M. S. Keetch, *Stability regions in the numerical treatment of Volterra integral equations*, SIAM J. Numer. Anal. 15 (1978), pp. 394–417.
- [4] — *Stability analysis of certain Runge-Kutta procedures for Volterra integral equations*, ACM Trans. Math. Software 4 (1978), pp. 305–315.
- [5] R. N. Castleton and L. J. Grimm, *A first order method for differential equations of neutral type*, Math. Comp. 27 (1973), pp. 571–577.
- [6] L. M. Delves and J. E. Walsh (eds.), *Numerical solution of integral equations*, Proceedings of Manchester-Liverpool Summer School, Clarendon Press, Oxford 1974.
- [7] L. Garey, *Solving nonlinear second kind Volterra equations by modified increment methods*, SIAM J. Numer. Anal. 12 (1975), pp. 501–508.
- [8] — *Block methods for nonlinear Volterra integral equations*, BIT 15 (1975), pp. 401–408.
- [9] B. Hall and J. M. Watt (eds.), *Modern numerical methods for ordinary differential equations*, Clarendon Press, Oxford 1976.
- [10] P. Henrici, *Discrete variable methods in ordinary differential equations*, John Wiley, New York, 1968.
- [11] Z. Jackiewicz, *The numerical solution of Volterra functional differential equations of neutral type*, SIAM J. Numer. Anal. 18 (1981), pp. 615–626.
- [12] — *One-step methods for the numerical solution of Volterra functional differential equations of neutral type*, Applicable Anal. 12 (1981), pp. 1–11.
- [13] — and M. Kwapisz, *On numerical integration of implicit ordinary differential equations*, Apl. Mat. 26 (1981), pp. 97–110.
- [14] M. Kuczma, *Functional equations in a single variable*, PWN, Warszawa 1968.
- [15] J. D. Lambert, *Computational methods in ordinary differential equations*, John Wiley, New York 1973.
- [16] B. Noble, *Instability when solving Volterra integral equations of the second kind by multistep methods*, pp. 23–39 in: *Conference on the Numerical Solution of Differential Equations*, Lecture Notes in Mathematics, Vol. 109, Springer-Verlag, Berlin-Heidelberg-New York 1969.
- [17] P. Pouzet, *Méthode d'intégration numérique des équations intégrales et intégrales différentielles du type de Volterra de seconde espèce, Formules de Runge-Kutta*, Symposium on the Numerical Treatment of Ordinary Differential Equations, Integral and Integro-Differential Equations (Rome, 1960), Birkhäuser Verlag, Basel 1960, pp. 362–368.
- [18] L. Tavernini, *Linear multistep methods for the numerical solution of Volterra functional differential equations*, Applicable Anal. 1 (1973), pp. 169–185.

- [19] T. S. Zverkina, *A modification of finite-difference methods for integrating ordinary differential equations with nonsmooth solutions* (in Russian), *Ž. Vyčisl. Mat. i Mat. Fiz.* (4) (1964), pp. 149–160.

DEPT. OF MATHEMATICAL SCIENCES
UNIVERSITY OF ARKANSAS
FAYETTEVILLE, AR 72701, USA

INSTITUTE OF MATHEMATICS
UNIVERSITY OF GDAŃSK
80-952 GDAŃSK

Received on 1985.09.05
