Electronic Transactions on Numerical Analysis.
Volume 26, pp. 190-208, 2007.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

EXTENSIONS OF THE HHT-α METHOD TO DIFFERENTIAL-ALGEBRAIC EQUATIONS IN MECHANICS*

LAURENT O. JAY† AND DAN NEGRUT‡

Abstract. We present second order extensions of the Hilber-Hughes-Taylor-α (HHT-α) method for systems of overdetermined differential-algebraic equations (ODAEs) arising, for example, in mechanics. A detailed analysis of extensions of the HHT-α method is given. In particular a local and global error analysis is presented. Second order convergence is theoretically demonstrated and practically illustrated by numerical experiments. A new variable stepsize formula is proposed which preserves the second order of the method.

Key words. differential-algebraic equations, HHT-α method, variable stepsize

AMS subject classifications. 65L05, 65L06, 65L80, 70F20, 70H45

1. Introduction.

The Hilber-Hughes-Taylor-α (HHT-α) method [6, 7] and its generalizations, such as the generalized-α method [3, 4], are widely used in structural and flexible multibody dynamics. This paper is concerned with extending the HHT-α method to systems of overdetermined differential-algebraic equations (ODAEs) with index 3 constraints and their underlying index 2 constraints, e.g., to systems in mechanics having holonomic constraints. An extension of the HHT-α method to index 2 DAEs, e.g., to systems in mechanics with nonholonomic constraints, is briefly discussed as well. We have found extensions of the HHT-α method preserving its second order convergence. Our extensions are indirect in the sense that we make use of the partitioned and additive structures of the ODAEs. Detailed mathematical proofs of second order convergence of extensions of the HHT-α method to the systems of ODAEs considered are given. A new variable stepsize formula is proposed which preserves the second order of the method. Second order convergence of these extensions is numerically illustrated on two test problems.

For DAEs global error estimates generally do not follow directly from local error estimates. The error propagation mechanism of a method for DAEs is usually more complicated than for ordinary differential equations (ODEs). In particular, for DAEs one cannot generally infer a global order of convergence directly from its local error estimates, as an order reduction may occur due to error propagation [1]. Analysis of the direct extension of the HHT-α method to linear DAEs was performed in [2]. It was shown that for semi-explicit index 3 linear DAEs the direct application of the HHT-α method is inconsistent and suffers from instabilities, but that it may still converge when applied with constant stepsize, similarly to BDF methods [1]. A first order extension of the HHT-α method to holonomically constrained mechanical systems was proposed in [12] and is based on projecting the solution of the underlying ODEs onto the constraints after each step. In [11] the direct application of the HHT-α method to index 3 holonomically constrained mechanical systems is considered, but no convergence result is given. The extensions of the HHT-α method that we present in this paper have second order convergence without relying on underlying ODEs and they also directly preserve the underlying index 2 constraints.

This paper is organized as follows. In Section 2 we describe the original HHT-α method for second order systems of ODEs and we propose a new variable stepsize formula preserving

*Received June 28, 2006. Accepted for publication December 21, 2006. Recommended by V. Mehrmann. This material is based upon work supported by the National Science Foundation under Grant No. 9983708.

†Department of Mathematics, 14 MacLean Hall, The University of Iowa, Iowa City, IA 52242-1419 (ljay@math.uiowa.edu and na.ljay@na-net.ornl.gov).

‡Department of Mechanical Engineering, University of Wisconsin-Madison, 3142 Engineering Centers Bldg, 1550 Engineering Dr., Madison, WI 53706-1572 (negrut@engr.wisc.edu).

ETNA
Kent State University
etna@mcs.kent.edu

the second order of the method. In Section 3 we present extensions of the HHT-α method to systems of ODAEs with index 3 constraints and underlying index 2 constraints. In Section 4 we give a detailed analysis of extensions of the HHT-α method to these systems of ODAEs. In particular, we show existence and uniqueness of the numerical solution, we analyze the local error of the method, its stability with respect to consistent perturbations in the initial values, and prove its global second order convergence. In Section 5 we illustrate the second order convergence of an extended HHT-α method on two test problems. In Section 6 we propose an extension of the HHT-α method for index 2 DAEs. A short conclusion is given in Section 7.

2. The HHT-α method for second order systems of ODEs. Second order systems of ODEs $y'' = f(t, y, y')$ are equivalent to

$$y' = z, \quad z' = a, \quad a = f(t, y, z). \tag{2.1}$$

In mechanics, y represents generalized coordinates, z represents the corresponding velocities, a represents the corresponding accelerations, $f(t, y, z) = M^{-1}F(t, y, z)$ where M is the mass matrix and $F(t, y, z)$ represents external forces. The HHT-α method for the system of equations (2.1) can be expressed as an implicit non-standard one-step method

$$(y_1, z_1, a_1) = \Phi_h(y_0, z_0, a_0)$$

as follows [6, 7]

$$y_1 = y_0 + hz_0 + \frac{h^2}{2}\left((1-2\beta)a_0 + 2\beta a_1\right), \tag{2.2a}$$

$$z_1 = z_0 + h\left((1-\gamma)a_0 + \gamma a_1\right), \tag{2.2b}$$

$$a_1 = (1+\alpha)f(t_1, y_1, z_1) - \alpha f(t_0, y_0, z_0), \tag{2.2c}$$

where h is the stepsize and $t_1 := t_0 + h$. For the HHT-α method (2.2) the coefficients α, β, γ are chosen according to

$$\alpha \in \left[-\frac{1}{3}, 0\right], \qquad \beta = \frac{(1-\alpha)^2}{4}, \qquad \gamma = \frac{1}{2} - \alpha.$$

The free coefficient α is a damping parameter. Notice that the notation for a_0 and a_1 may be misleading. These values are not really approximations of $a(t_0)$ and $a(t_1)$ respectively, but of $a(t_0 + \alpha h)$ and $a(t_1 + \alpha h)$. The coefficient $\gamma = \frac{1}{2} - \alpha$ is determined such that the method is of local order 2 in z when $a_0 - a(t_0 + \alpha h) = O(h^2)$. When $a_0 - a(t_0 + \alpha h) = O(h)$, e.g., when $a_0 = a(t_0) = f(t_0, y_0, z_0)$ the method is only of local order 1 in z for the first step, i.e., $z_1 - z(t_0 + h) = O(h^2)$. However, in this situation since $a_1 - a(t_1 + \alpha h) = O(h^2)$ the next step (y_2, z_2, a_2) has nevertheless an error estimate in z of the form $z_2 - z(t_1 + h) = O(h^3)$. The HHT-$\alpha$ method is thus self-correcting, explaining in part its global second order convergence even when a_0 is taken as $a_0 = f(t_0, y_0, z_0)$. The HHT-α method is generally applied with constant stepsize in order to keep its second order accuracy. For $a_0(\widetilde{h})$ coming from the previous step with stepsize $\widetilde{h}$, by changing the stepsize from $\widetilde{h}$ to $h \neq \widetilde{h}$, a_1 is no more an approximation of local order 1 to $a(t_1 + \alpha h)$, i.e., it does not satisfy $a_1 - a(t_1 + \alpha h) = O(h^2)$. Hence, without any modification the HHT-α method reduces to a first order method for all variables. To reestablish global second order convergence while still allowing stepsize changes using $a_0(\widetilde{h})$ from the previous step taken with stepsize $\widetilde{h}$, one can replace the definition of a_0 for the current step by

$$a_0 := f(t_0, y_0, z_0) + \frac{h}{\widetilde{h}}(a_0(\widetilde{h}) - f(t_0, y_0, z_0)). \tag{2.3}$$

ETNA
Kent State University
etna@mcs.kent.edu

Simply taking $a_0 := \alpha f(t_1, y_1, z_1) + (1-\alpha) f(t_0, y_0, z_0)$ analogously to (2.2c) leads to the expression $z_1 = z_0 + h\left(\frac{1}{2} f(t_0, y_0, z_0) + \frac{1}{2} f(t_1, y_1, z_1)\right)$ which corresponds to the trapezoidal rule which has no damping parameter α and is thus not recommended.

3. Extensions of the HHT-α method to ODAEs. We consider semi-explicit index 3 DAEs of the form

$$y' = z, \tag{3.1a}$$
$$z' = a + r(t, y, \lambda), \tag{3.1b}$$
$$a = f(t, y, z), \tag{3.1c}$$
$$0 = g(t, y), \tag{3.1d}$$

where we assume that $g_y(t, y) r_\lambda(t, y, \lambda)$ is invertible in the region of interest. In mechanics (3.1d) represents holonomic constraints, λ represents Lagrange multipliers, and $r(t, y, \lambda) = -M^{-1} g_y^T(t, y)\lambda$ where M is the mass matrix and $-g_y^T(t, y)\lambda$ represents reaction forces coming from the constraints [1]. Differentiating (3.1d) once with respect to t, we obtain additional constraints

$$0 = g_t(t, y) + g_y(t, y) z. \tag{3.1e}$$

The whole system of ODAEs (3.1) is of index 2. One more differentiation of (3.1e) with respect to t leads to

$$0 = g_{tt}(t, y) + 2g_{ty}(t, y) z + g_{yy}(t, y)(z, z) + g_y(t, y)(f(t, y, z) + r(t, y, \lambda)). \tag{3.2}$$

We will not make a direct use of these additional constraints (3.2) in the numerical scheme (3.4) below. Nevertheless, it will be useful to consider them in the analysis of the method. From the constraints (3.2), one more differentiation gives an expression for λ',

$$\begin{aligned} \lambda' = (-g_y r_\lambda)^{-1} \Big(& g_{ttt} + 3g_{tty} z + 3g_{tyy}(z, z) + g_{yyy}(z, z, z) + 3g_{ty}(f + r) \\ & + 3g_{yy}(z, f + r) + g_y(f_t + f_y z + f_z(f + r) + r_t + r_y z)\Big) \end{aligned} \tag{3.3}$$

where we have not written explicitly the arguments (t, y, z, λ) for f, r, g, f_y, r_y, g_y, etc.

We propose a new generalization of the HHT-α method for the system (3.1). Although different in essence our approach is reminiscent of the GGL/stabilized index 2 formulation [5]. Here, instead of artificially introducing additional new algebraic variables in (3.1a), we consider directly the systems of ODAEs (3.1). Given (y_0, z_0, a_0) we define the extended HHT-α method for (3.1) as follows

$$y_1 = y_0 + h z_0 + \frac{h^2}{2}\left((1 - 2\beta) a_0 + 2\beta a_1\right) + \frac{h^2}{2}\left((1 - b) R_0 + b R_1\right), \tag{3.4a}$$
$$z_1 = z_0 + h\left((1 - \gamma) a_0 + \gamma a_1\right) + \frac{h}{2}\left(R_0 + R_1\right), \tag{3.4b}$$
$$a_1 = (1 + \alpha) f(t_1, y_1, z_1) - \alpha f(t_0, y_0, z_0), \tag{3.4c}$$

where $b \neq 1/2$ is a free coefficient,

$$R_0 := r(t_0, y_0, \Lambda_0), \qquad R_1 := r(t_1, y_1, \Lambda_1), \tag{3.4d}$$

and Λ_0 is not a value λ_0 coming from the previous step, but Λ_0 and Λ_1 are locally determined by the two sets of constraints

$$0 = g(t_1, y_1), \qquad 0 = g_t(t_1, y_1) + g_y(t_1, y_1) z_1. \tag{3.4e}$$

This determination of Λ_0 and Λ_1 is an important point. The numerical solution (y_1, z_1) thus satisfies both constraints (3.1d)-(3.1e) at each timestep. We propose the simple choice $b = 0$. For $b = 0$ and $\alpha = 0$ the method is an additive combination of the 2-stage Lobatto IIIA and Lobatto IIIB implicit Runge-Kutta coefficients, and is known to be of second order for all variables [9] (note that it is not the combination of Lobatto IIIA and Lobatto IIIB coefficients given in [8] since for unconstrained problems the HHT-α method is simply equivalent to the trapezoidal rule, the 2-stage Lobatto IIIA method). To make the method less implicit, one can replace R_1 in (3.4d) by $R_1 := r(t_1, \widetilde{y}_1, \Lambda_1)$ where

$$\widetilde{y}_1 := y_0 + hz_0. \tag{3.5}$$

Another possibility is to take $b = 0$ and to replace the expression $(R_0 + R_1)/2$ in (3.4b) by the midpoint approximation

$$R_{1/2} = r\left(t_0 + \frac{h}{2}, \frac{y_0 + y_1}{2}, \Lambda_{1/2}\right)$$

and also with y_1 replaced by (3.5). The results given in this paper remain valid with these simplifications under some minor modifications. In particular, second order global convergence as shown in Theorem 4.5 also holds.

4. Analysis of the extended HHT-α method for ODAEs. First we show existence and uniqueness of the numerical solution of the extended HHT-α method (3.4).

THEOREM 4.1. *Consider the overdetermined system of DAEs (3.1) with initial conditions $(y_0, z_0, a_0) = (y_0(h), z_0(h), a_0(h))$ depending on h and satisfying*

$$g(t_0, y_0) = O(h^3), \quad g_t(t_0, y_0) + g_y(t_0, y_0)z_0 = O(h^2), \quad a_0 - a(t_0 + \alpha h) = O(h).$$

Then for $0 \le h \le h_0$ there exists a unique solution $(y_1, z_1, a_1, \Lambda_0, \Lambda_1)$ depending on h to the system of equations (3.4) in a neighborhood of $(y_0, z_0, a_0, \lambda_0, \lambda_0)$ where $\lambda_0 = \lambda_0(h)$ satisfies

$$\begin{aligned} g_{tt}(t_0, y_0) + 2g_{ty}(t_0, y_0)z_0 + g_{yy}(t_0, y_0)(z_0, z_0) \\ + g_y(t_0, y_0)(f(t_0, y_0, z_0) + r(t_0, y_0, \lambda_0)) = O(h). \end{aligned}$$

Moreover, we have the estimates

$$y_1 - y_0 = O(h), \quad z_1 - z_0 = O(h), \quad a_1 - a_0 = O(h), \tag{4.1a}$$
$$\Lambda_0 - \lambda_0 = O(h), \quad \Lambda_1 - \lambda_0 = O(h). \tag{4.1b}$$

REMARK 4.2. Note that the numerical solution $(y_1, z_1, a_1, \Lambda_0, \Lambda_1)$ is functionnally independent of λ_0. The value λ_0 only indicates a solution branch to which the numerical solution is close. Varying slightly λ_0 to $\lambda_0 + \delta$ with a small perturbation $\delta = o(1)$ does not change the numerical solution $(y_1, z_1, a_1, \Lambda_0, \Lambda_1)$.

Proof. The proof of this theorem can be made by application of the implicit function theorem. We first introduce directly the definition (3.4d) of R_0, R_1 into (3.4a) and (3.4b). Then we also replace partially some expressions for y_1 and z_1 explicitly in (3.4e). Multiplying the two equations of (3.4e) by $2/h^2$ and $1/h$ respectively, we obtain the equivalent system of equations

$$0 = y_1 - \left(y_0 + hz_0 + \frac{h^2}{2}\big((1-2\beta)a_0 + 2\beta a_1 + (1-b)r(t_0, y_0, \Lambda_0) + br(t_1, y_1, \Lambda_1)\big)\right),$$

ETNA
Kent State University
etna@mcs.kent.edu

$$
\begin{aligned}
0 &= z_1 - \Big(z_0 + h\big((1-\gamma)a_0 + \gamma a_1 + \frac{1}{2}r(t_0,y_0,\Lambda_0) + \frac{1}{2}r(t_1,y_1,\Lambda_1)\big)\Big),\\
0 &= a_1 - \Big((1+\alpha)f(t_1,y_1,z_1) - \alpha f(t_0,y_0,z_0)\Big),\\
0 &= \frac{2}{h^2}g\Big(t_1, y_0 + hz_0 + \frac{h^2}{2}\big((1-2\beta)a_0 + 2\beta a_1 + (1-b)r(t_0,y_0,\Lambda_0) + br(t_1,y_1,\Lambda_1)\big)\Big),\\
0 &= \frac{1}{h}g_t(t_1,y_1)\\
&\quad + \frac{1}{h}g_y(t_1,y_1)\left(z_0 + h\big((1-\gamma)a_0 + \gamma a_1 + \frac{1}{2}r(t_0,y_0,\Lambda_0) + \frac{1}{2}r(t_1,y_1,\Lambda_1)\big)\right).
\end{aligned}
$$

Replacing a_1 by its expression (3.4c) in the last two equations and then expanding in h around (t_0, y_0, z_0) we obtain

$$
\begin{aligned}
0 &= \frac{2}{h^2}g(t_0,y_0) + \frac{2}{h}(g_t(t_0,y_0) + g_y(t_0,y_0)z_0)\\
&\quad + g_y(t_0,y_0)\Big((1-2\beta)a_0 + 2\beta f(t_0,y_0,z_0) + (1-b)r(t_0,y_0,\Lambda_0) + br(t_0,y_0,\Lambda_1)\Big)\\
&\quad + g_{tt}(t_0,y_0) + 2g_{ty}(t_0,y_0)z_0 + g_{yy}(t_0,y_0)(z_0,z_0) + O(h),\\
0 &= \frac{1}{h}(g_t(t_0,y_0) + g_y(t_0,y_0)z_0) + g_{tt}(t_0,y_0) + 2g_{ty}(t_0,y_0)z_0 + g_{yy}(t_0,y_0)(z_0,z_0)\\
&\quad + g_y(t_0,y_0)\Big((1-\gamma)a_0 + \gamma f(t_0,y_0,z_0) + \frac{1}{2}r(t_0,y_0,\Lambda_0) + \frac{1}{2}r(t_0,y_0,\Lambda_1)\Big) + O(h).
\end{aligned}
$$

By using the hypotheses of the theorem all equations are satisfied at $h = 0$ by

$$(y_1(0), z_1(0), a_1(0), \Lambda_0(0), \Lambda_1(0)) := (y_0(0), z_0(0), a_0(0), \lambda_0(0), \lambda_0(0)).$$

The Jacobian at $h = 0$ of the above equations with respect to $(y_1, z_1, a_1, \Lambda_0, \Lambda_1)$ is given by

$$
\begin{bmatrix}
I & O & O & O & O\\
O & I & O & O & O\\
* & * & I & O & O\\
* & * & * & (1-b)M_0 & bM_0\\
* & * & * & \frac{1}{2}M_0 & \frac{1}{2}M_0
\end{bmatrix}
$$

where $M_0 := g_y(t_0,y_0)r_\lambda(t_0,y_0,\lambda_0)$ is invertible. Since

$$
\begin{bmatrix}
(1-b)M_0 & bM_0\\
\frac{1}{2}M_0 & \frac{1}{2}M_0
\end{bmatrix}
=
\begin{bmatrix}
(1-b) & b\\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix}
\otimes M_0.
$$

the Jacobian at $h = 0$ is invertible provided $b \neq 1/2$. The conclusion and the estimates (4.1) now follow by application of the implicit function theorem. □

We now consider local error estimates:

THEOREM 4.3. *Consider the overdetermined system of DAEs (3.1) with initial conditions* (y_0, z_0, a_0) *at* t_0 *satisfying*

$$g(t_0,y_0) = 0, \quad g_t(t_0,y_0) + g_y(t_0,y_0)z_0 = 0, \quad a_0 - a(t_0 + \alpha h) = O(h),$$

and let λ_0 *be such that*

$$g_{tt}(t_0,y_0) + 2g_{ty}(t_0,y_0)z_0 + g_{yy}(t_0,y_0)(z_0,z_0) + g_y(t_0,y_0)(f(t_0,y_0,z_0) + r(t_0,y_0,\lambda_0)) = 0.$$

Then for $0 \le h \le h_0$ *the numerical solution* (y_1, z_1, a_1) *at* $t_1 = t_0 + h$ *to the system of equations (3.4) satisfies*

$$y_1 - y(t_1) = O(h^3), \quad z_1 - z(t_1) = O(h^2), \quad a_1 - a(t_1 + \alpha h) = O(h^2), \tag{4.2}$$

where $(y(t), z(t))$ *is the exact solution to (3.1) at* t *passing through* (y_0, z_0) *at* t_0*. If in addition we assume that*

$$a_0 - a(t_0 + \alpha h) = O(h^2) \tag{4.3}$$

then

$$z_1 - z(t_1) = O(h^3). \tag{4.4}$$

Proof. The Taylor series of the exact solution $(y(t), z(t))$ at $t_1 = t_0 + h$ satisfies

$$y(t_1) = y_0 + hz_0 + \frac{h^2}{2}(f_0 + r_0) + O(h^3), \qquad z(t_1) = z_0 + h(f_0 + r_0) + O(h^2),$$

where $f_0 := f(t_0, y_0, z_0)$ and $r_0 := r(t_0, y_0, \lambda_0)$. We know from Theorem 4.1 that $\Lambda_0(h) - \lambda_0 = O(h)$ and $\Lambda_1(h) - \lambda_0 = O(h)$. For the numerical solution (y_1, z_1) we have by direct application of the estimates (4.1) in the definition (3.4a)-(3.4b)

$$y_1 = y_0 + hz_0 + \frac{h^2}{2}(f_0 + r_0) + O(h^3), \qquad z_1 = z_0 + h(f_0 + r_0) + O(h^2).$$

Hence, we obtain $y_1 - y(t_1) = O(h^3)$ and $z_1 - z(t_1) = O(h^2)$. From (2.1) and (3.4c) for a_1 we have

$$a(t_1 + \alpha h) = f_0 + (1 + \alpha)h(f_{t0} + f_{y0}z_0 + f_{z0}(f_0 + r_0)) + O(h^2)$$

and

$$a_1 = f_0 + (1 + \alpha)h(f_{t0} + f_{y0}z_0 + f_{z0}(f_0 + r_0)) + O(h^2).$$

A direct consequence is the estimate $a_1 - a(t_1 + \alpha h) = O(h^2)$. It remains to show (4.4) when (4.3) holds. The condition $a_0 - a(t_0 + \alpha h) = O(h^2)$ is equivalent to

$$a_0 = f_0 + h\alpha(f_{t0} + f_{y0}z_0 + f_{z0}(f_0 + r_0)) + O(h^2).$$

The Taylor series of $z(t_1)$ at t_0 satisfies

$$z(t_1) = z_0 + h(f_0 + r_0) + \frac{h^2}{2}\left(f_{t0} + f_{y0}z_0 + f_{z0}(f_0 + r_0) + r_{t0} + r_{y0}z_0 + r_{\lambda 0}\lambda_0'\right) + O(h^3)$$

where λ_0' corresponds to the expression (3.3) evaluated at $(t_0, y_0, z_0, \lambda_0)$. Since $\gamma = 1/2 - \alpha$ we get

$$\left(\frac{1}{2} + \alpha\right)a_0 + \left(\frac{1}{2} - \alpha\right)a_1 = f_0 + \frac{h}{2}(f_{t0} + f_{y0}z_0 + f_{z0}(f_0 + r_0)) + O(h^2).$$

We also have

$$R_0 = r_0 + hr_{\lambda 0}\Lambda_0'(0) + O(h^2), \qquad R_1 = r_0 + h(r_{t0} + r_{y0}z_0 + r_{\lambda 0}\Lambda_1'(0)) + O(h^2).$$

ETNA
Kent State University
etna@mcs.kent.edu

Putting all previous estimates together we obtain

$$z_1 = z_0 + h(f_0 + r_0) + \frac{h^2}{2}\Big(f_{t0} + f_{y0}z_0 + f_{z0}(f_0 + r_0) + r_{t0} + r_{y0}z_0 + r_{\lambda 0}(\Lambda_0'(0) + \Lambda_1'(0))\Big) + O(h^3).$$

Thus, it remains to show that $\Lambda_0'(0) + \Lambda_1'(0) = \lambda_0'$. This expression can be obtained by expanding

$$0 = \frac{1}{h^2}\left(g_t(t_1, y_1) + g_y(t_1, y_1)z_1\right) \tag{4.5}$$

around $(t_0, y_0, z_0, \lambda_0)$ into h-powers and then letting $h \to 0$. First, we write $y_1 = y_0 + \delta$ with $\delta := hz_0 + h^2/2(f_0 + r_0) + O(h^3) = O(h)$ and we expand in h and δ

$$g_t(t_0 + h, y_0 + \delta) = g_{t0} + g_{tt0}h + g_{ty0}\delta + \frac{1}{2}\left(g_{ttt0}h^2 + 2g_{tty0}h\delta + g_{tyy0}(\delta, \delta)\right) + O(h^3),$$

$$g_y(t_0 + h, y_0 + \delta) = g_{y0} + g_{ty0}h + g_{yy0}\delta + \frac{1}{2}\left(g_{tty0}h^2 + 2g_{tyy0}h\delta + g_{yyy0}(\delta, \delta)\right) + O(h^3).$$

Expanding (4.5) in h-powers, grouping the terms, and letting $h \to 0$ we finally obtain

$$\begin{aligned} 0 = \; & g_{ttt0} + 3g_{tty0}z_0 + 3g_{tyy0}(z_0, z_0) + g_{yyy0}(z_0, z_0, z_0) + 3g_{ty0}(f_0 + r_0) \\ & + 3g_{yy0}(z_0, f_0 + r_0) + g_{y0}(f_{t0} + f_{y0}z_0 + f_{z0}(f_0 + r_0) + r_{t0} + r_{y0}z_0) \\ & + g_{y0}r_{\lambda 0}(\Lambda_0'(0) + \Lambda_1'(0)). \end{aligned}$$

From (3.3) this leads to the desired result $\Lambda_0'(0) + \Lambda_1'(0) = \lambda_0'$ and therefore $z_1 - z(t_1) = O(h^3)$. □

To analyze the error propagation we introduce the projectors

$$Q(t, y, \lambda) := r_\lambda(t, y, \lambda)(g_y(t, y)r_\lambda(t, y, \lambda))^{-1}g_y(t, y), \qquad P(t, y, \lambda) := I - Q(t, y, \lambda).$$

They have the properties

$$\begin{aligned} &Q(t, y, \lambda)r_\lambda(t, y, \lambda) = r_\lambda(t, y, \lambda), \qquad && g_y(t, y)Q(t, y, \lambda) = g_y(t, y), \\ &P(t, y, \lambda)r_\lambda(t, y, \lambda) = 0, && g_y(t, y)P(t, y, \lambda) = 0. \end{aligned}$$

Before proving global convergence, we need to study changes in the numerical solution with respect to perturbations in consistent initial conditions:

THEOREM 4.4. *Consider $(\widetilde{y}_k, \widetilde{z}_k, \widetilde{a}_k)$ $(\widehat{y}_k, \widehat{z}_k, \widehat{a}_k)$ at t_k satisfying the constraints (3.1d) and (3.1e). Let $\Delta y_k := \widehat{y}_k - \widetilde{y}_k$, $\Delta z_k := \widehat{z}_k - \widetilde{z}_k$, $\Delta a_k := \widehat{a}_k - \widetilde{a}_k$, satisfying*

$$\Delta y_k = O(h), \quad \Delta z_k = O(h), \quad \Delta a_k = O(h),$$

and let $(\widetilde{y}_{k+1}, \widetilde{z}_{k+1}, \widetilde{a}_{k+1})$ and $(\widehat{y}_{k+1}, \widehat{z}_{k+1}, \widehat{a}_{k+1})$ be the corresponding HHT-α solutions (3.4). Then we have

$$\begin{aligned} P_{k+1}\Delta y_{k+1} &= P_k\Delta y_k + hP_k\Delta z_k + (1/2 - \beta)\, h^2 P_k\Delta a_k \\ &\quad + O(h\|P_k\Delta y_k\| + h^2\|P_k\Delta z_k\| + h^3\|\Delta a_k\|), \\ hP_{k+1}\Delta z_{k+1} &= hP_k\Delta z_k + (1 - \gamma)\, h^2 P_k\Delta a_k \\ &\quad + O(h^2\|P_k\Delta y_k\| + h^2\|P_k\Delta z_k\| + h^3\|\Delta a_k\|), \\ h^2\Delta a_{k+1} &= O(h^2\|P_k\Delta y_k\| + h^2\|P_k\Delta z_k\| + h^3\|\Delta a_k\|), \\ Q_{k+1}\Delta y_{k+1} &= O(\|P_{k+1}\Delta y_{k+1}\|), \\ Q_{k+1}\Delta z_{k+1} &= O(\|P_{k+1}\Delta y_{k+1}\| + \|P_{k+1}\Delta z_{k+1}\|), \end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

where $P_k := P(t_k, \widetilde{y}_k, \widetilde{\lambda}_k)$, $P_{k+1} := P(t_{k+1}, \widetilde{y}_{k+1}, \widetilde{\lambda}_{k+1})$, *and* $\widetilde{\lambda}_k, \widetilde{\lambda}_{k+1}$ *are such that the constraints (3.2) are satisfied for* $(\widetilde{y}_k, \widetilde{z}_k, \widetilde{\lambda}_k)$ *and* $(\widetilde{y}_{k+1}, \widetilde{z}_{k+1}, \widetilde{\lambda}_{k+1})$ *respectively.*

Proof. Let $\widehat{\lambda}_k$ be such that the constraints (3.2) are satisfied for $(\widehat{y}_k, \widehat{z}_k, \widehat{\lambda}_k)$ and satisfying $\widehat{\lambda}_k - \widetilde{\lambda}_k = O(h)$. By Theorem 4.1 we have

$$\begin{aligned}
&\widetilde{y}_{k+1} - \widetilde{y}_k = O(h), \quad \widetilde{z}_{k+1} - \widetilde{z}_k = O(h), \quad \widetilde{a}_{k+1} - \widetilde{a}_k = O(h),\\
&\widetilde{\Lambda}_{k0} - \widetilde{\lambda}_k = O(h), \quad \widetilde{\Lambda}_{k1} - \widetilde{\lambda}_k = O(h),\\
&\widehat{y}_{k+1} - \widehat{y}_k = O(h), \quad \widehat{z}_{k+1} - \widehat{z}_k = O(h), \quad \widehat{a}_{k+1} - \widehat{a}_k = O(h),\\
&\widehat{\Lambda}_{k0} - \widehat{\lambda}_k = O(h), \quad \widehat{\Lambda}_{k1} - \widehat{\lambda}_k = O(h).
\end{aligned}$$

Hence, we also have

$$\Delta y_{k+1} = O(h), \quad \Delta z_{k+1} = O(h), \quad \Delta a_{k+1} = O(h), \quad \Delta\Lambda_{k0} = O(h), \quad \Delta\Lambda_{k1} = O(h),$$

where $\Delta\Lambda_{k0} := \widehat{\Lambda}_{k0} - \widetilde{\Lambda}_{k0}$, and $\Delta\Lambda_{k1} := \widehat{\Lambda}_{k1} - \widetilde{\Lambda}_{k1}$. Substracting (3.4abc) for $(\widehat{y}_{k+1}, \widehat{z}_{k+1}, \widehat{a}_{k+1})$ from (3.4abc) for $(\widetilde{y}_{k+1}, \widetilde{z}_{k+1}, \widetilde{a}_{k+1})$ and linearizing around $(t_k, \widetilde{y}_k, \widetilde{z}_k, \widetilde{a}_k)$ we obtain

$$\begin{aligned}
\text{(4.6a)}\quad \Delta y_{k+1} = {}& \Delta y_k + h\Delta z_k + \frac{h^2}{2}\left((1-2\beta)\Delta a_k + 2\beta\Delta a_{k+1}\right)\\
&+\frac{h^2}{2}\left((1-b)(r_{y,k}\Delta y_k + r_{\lambda,k}\Delta\Lambda_{k0}) + b(r_{y,k+1}\Delta y_{k+1} + r_{\lambda,k+1}\Delta\Lambda_{k1})\right)\\
&+O(h^2\|\Delta y_k\|^2 + h^2\|\Delta y_{k+1}\|^2 + h^2\|\Delta\Lambda_{k0}\|^2 + h^2\|\Delta\Lambda_{k1}\|^2),\\
\text{(4.6b)}\quad \Delta z_{k+1} = {}& \Delta z_k + h\left((1-\gamma)\Delta a_k + \gamma\Delta a_{k+1}\right)\\
&+\frac{h}{2}\left(r_{y,k}\Delta y_k + r_{\lambda,k}\Delta\Lambda_{k0} + r_{y,k+1}\Delta y_{k+1} + r_{\lambda,k+1}\Delta\Lambda_{k1}\right)\\
&+O(h\|\Delta y_k\|^2 + h\|\Delta y_{k+1}\|^2 + h\|\Delta\Lambda_{k0}\|^2 + h\|\Delta\Lambda_{k1}\|^2),\\
\text{(4.6c)}\quad \Delta a_{k+1} = {}& (1+\alpha)\left(f_{y,k+1}\Delta y_{k+1} + f_{z,k+1}\Delta z_{k+1}\right) - \alpha\left(f_{y,k}\Delta y_k + f_{z,k}\Delta z_k\right)\\
&+O(\|\Delta y_k\|^2 + \|\Delta y_{k+1}\|^2 + \|\Delta z_k\|^2 + \|\Delta z_{k+1}\|^2)\\
={}& O(\|\Delta y_k\| + \|\Delta y_{k+1}\| + \|\Delta z_k\| + \|\Delta z_{k+1}\|).
\end{aligned}$$

From

$$0 = g(t_{k+1}, \widehat{y}_{k+1}), \qquad 0 = g(t_{k+1}, \widetilde{y}_{k+1}),$$

we obtain by linearization around $(t_{k+1}, \widetilde{y}_{k+1})$

$$\text{(4.7)}\qquad 0 = g_{y,k+1}\Delta y_{k+1} + O(\|\Delta y_{k+1}\|^2).$$

Introducing the expression (4.6a) for Δy_{k+1} we get

$$\begin{aligned}
\text{(4.8)}\qquad & -\frac{1}{2}\left((1-b)g_{y,k+1}r_{\lambda,k}h^2\Delta\Lambda_{k0} + bg_{y,k+1}r_{\lambda,k+1}h^2\Delta\Lambda_{k1}\right)\\
&= g_{y,k+1}\Delta y_k + hg_{y,k+1}\Delta z_k\\
&\quad + \frac{h^2}{2}\left((1-2\beta)g_{y,k+1}\Delta a_k + 2\beta g_{y,k+1}\Delta a_{k+1}\right)\\
&\quad + \frac{h^2}{2}\left((1-b)g_{y,k+1}r_{y,k}\Delta y_k + bg_{y,k+1}r_{y,k+1}\Delta y_{k+1}\right)\\
&\quad + O(h^2\|\Delta y_k\|^2 + \|\Delta y_{k+1}\|^2 + h^2\|\Delta\Lambda_{k0}\|^2 + h^2\|\Delta\Lambda_{k1}\|^2).
\end{aligned}$$

ETNA
Kent State University
etna@mcs.kent.edu

From

$$0 = g_t(t_{k+1}, \widehat{y}_{k+1}) + g_y(t_{k+1}, \widehat{y}_{k+1})\widehat{z}_{k+1}, \quad 0 = g_t(t_{k+1}, \widetilde{y}_{k+1}) + g_y(t_{k+1}, \widetilde{y}_{k+1})\widetilde{z}_{k+1},$$

we obtain by linearization around $(t_{k+1}, \widetilde{y}_{k+1}, \widetilde{z}_{k+1})$

$$\begin{aligned} 0 = {} & g_{ty,k+1}\Delta y_{k+1} + g_{yy,k+1}(\widetilde{z}_{k+1}, \Delta y_{k+1}) + g_{y,k+1}\Delta z_{k+1} \\ & + O(\|\Delta y_{k+1}\|^2 + \|\Delta z_{k+1}\|^2). \end{aligned} \tag{4.9}$$

Introducing the expression for $h\Delta z_{k+1}$ from (4.6b) we get

$$\begin{aligned} & -\frac{1}{2}\left(g_{y,k+1}r_{\lambda,k}h^2\Delta\Lambda_{k0} + g_{y,k+1}r_{\lambda,k+1}h^2\Delta\Lambda_{k1}\right) \\ & = hg_{ty,k+1}\Delta y_{k+1} + hg_{yy,k+1}(\widetilde{z}_{k+1}, \Delta y_{k+1}) + hg_{y,k+1}\Delta z_k \\ & \quad + h^2\left((1-\gamma)g_{y,k+1}\Delta a_k + \gamma g_{y,k+1}\Delta a_{k+1}\right) \\ & \quad + \frac{h^2}{2}\left(g_{y,k+1}r_{y,k}\Delta y_k + g_{y,k+1}r_{y,k+1}\Delta y_{k+1}\right) \\ & \quad + O(h^2\|\Delta y_k\|^2 + h\|\Delta y_{k+1}\|^2 + h\|\Delta z_{k+1}\|^2 + h^2\|\Delta\Lambda_{k0}\|^2 + h^2\|\Delta\Lambda_{k1}\|^2). \end{aligned} \tag{4.10}$$

Since by assumption $0 = g(t_k, \widehat{y}_k)$ and $0 = g(t_k, \widetilde{y}_k)$ we obtain by linearization around $(t_k, \widetilde{y}_k)$

$$0 = g_{y,k}\Delta y_k + O(\|\Delta y_k\|^2). \tag{4.11}$$

Therefore, we can estimate the term $g_{y,k+1}\Delta y_k$ in (4.8) by

$$g_{y,k+1}\Delta y_k = g_{y,k}\Delta y_k + O(h\|\Delta y_k\|) = O(h\|\Delta y_k\| + \|\Delta y_k\|^2).$$

We can also estimate the term $g_{y,k+1}\Delta z_k$ in (4.8) and (4.10) by

$$g_{y,k+1}\Delta z_k = g_{y,k}Q_k\Delta z_k + O(h\|\Delta z_k\|).$$

We also have in (4.8) and (4.10)

$$g_{y,k+1}\Delta a_k = g_{y,k}Q_k\Delta a_k + O(h\|\Delta a_k\|), \qquad g_{y,k+1}\Delta a_{k+1} = g_{y,k+1}Q_{k+1}\Delta a_{k+1}.$$

For $b \neq 1/2$ and h sufficiently small the matrix

$$\begin{bmatrix} (1-b)g_{y,k+1}r_{\lambda,k} & bg_{y,k+1}r_{\lambda,k+1} \\ g_{y,k+1}r_{\lambda,k} & g_{y,k+1}r_{\lambda,k+1} \end{bmatrix}$$

is invertible and has a bounded inverse. Therefore, from (4.8) and (4.10) we obtain the following estimates for $h^2\|\Delta\Lambda_{k0}\|$ and $h^2\|\Delta\Lambda_{k1}\|$,

$$\begin{aligned} h^2\|\Delta\Lambda_{k0}\| = O\Big(& h\|\Delta y_k\| + \|\Delta y_k\|^2 + h\|\Delta y_{k+1}\| + \|\Delta y_{k+1}\|^2 + h\|Q_k\Delta z_k\| \\ & + h^2\|\Delta z_k\| + h\|\Delta z_{k+1}\|^2 + h^2\|Q_k\Delta a_k\| + h^3\|P_k\Delta a_k\| \\ & + h^2\|Q_{k+1}\Delta a_{k+1}\| + h^2\|\Delta\Lambda_{k0}\|^2 + h^2\|\Delta\Lambda_{k1}\|^2\Big), \end{aligned} \tag{4.12a}$$

$$\begin{aligned} h^2\|\Delta\Lambda_{k1}\| = O\Big(& h\|\Delta y_k\| + \|\Delta y_k\|^2 + h\|\Delta y_{k+1}\| + \|\Delta y_{k+1}\|^2 + h\|Q_k\Delta z_k\| \\ & + h^2\|\Delta z_k\| + h\|\Delta z_{k+1}\|^2 + h^2\|Q_k\Delta a_k\| + h^3\|P_k\Delta a_k\| \\ & + h^2\|Q_{k+1}\Delta a_{k+1}\| + h^2\|\Delta\Lambda_{k0}\|^2 + h^2\|\Delta\Lambda_{k1}\|^2\Big). \end{aligned} \tag{4.12b}$$

Multiplying (4.6a) and (4.6b) from the left by P_{k+1} and hP_{k+1} respectively, using $P_{k+1}r_{\lambda,k} = O(h), P_{k+1}r_{\lambda,k+1} = 0$, we get

$$\begin{aligned}
(4.13\text{a})\quad P_{k+1}\Delta y_{k+1} &= P_k\Delta y_k + hP_k\Delta z_k + \frac{h^2}{2}\left((1-2\beta)P_k\Delta a_k + 2\beta P_{k+1}\Delta a_{k+1}\right)\\
&\quad + O(h\|\Delta y_k\| + h^2\|\Delta y_{k+1}\| + h^2\|\Delta z_k\| + h^3\|\Delta a_k\| + h^3\|\Delta\Lambda_{k0}\|\\
&\quad + h^2\|\Delta y_k\|^2 + h^2\|\Delta y_{k+1}\|^2 + h^2\|\Delta\Lambda_{k0}\|^2 + h^2\|\Delta\Lambda_{k1}\|^2),\\
(4.13\text{b})\quad hP_{k+1}\Delta z_{k+1} &= hP_k\Delta z_k + h^2\left((1-\gamma)P_k\Delta a_k + \gamma P_{k+1}\Delta a_{k+1}\right)\\
&\quad + O(h^2\|\Delta y_k\| + h^2\|\Delta y_{k+1}\| + h^2\|\Delta z_k\| + h^3\|\Delta a_k\| + h^3\|\Delta\Lambda_{k0}\|\\
&\quad + h^2\|\Delta y_k\|^2 + h^2\|\Delta y_{k+1}\|^2 + h^2\|\Delta\Lambda_{k0}\|^2 + h^2\|\Delta\Lambda_{k1}\|^2).
\end{aligned}$$

Inserting the estimates (4.12) and (4.6c) into (4.13) we obtain

$$\begin{aligned}
(4.14\text{a})\quad P_{k+1}\Delta y_{k+1} &= P_k\Delta y_k + hP_k\Delta z_k + (1/2-\beta)h^2P_k\Delta a_k\\
&\quad + O(h\|\Delta y_k\| + h^2\|\Delta y_{k+1}\| + h^2\|\Delta z_k\| + h^2\|\Delta z_{k+1}\| + h^3\|\Delta a_k\|),\\
(4.14\text{b})\quad hP_{k+1}\Delta z_{k+1} &= hP_k\Delta z_k + h^2(1-\gamma)P_k\Delta a_k\\
&\quad + O(h^2\|\Delta y_k\| + h^2\|\Delta y_{k+1}\| + h^2\|\Delta z_k\| + h^2\|\Delta z_{k+1}\| + h^3\|\Delta a_k\|).
\end{aligned}$$

From (4.7) we have

$$Q_{k+1}\Delta y_{k+1} = O(\|\Delta y_{k+1}\|^2).$$

Thus,

$$\begin{aligned}
\Delta y_{k+1} &= P_{k+1}\Delta y_{k+1} + Q_{k+1}\Delta y_{k+1} = P_{k+1}\Delta y_{k+1} + O(\|\Delta y_{k+1}\|^2)\\
(4.15)\qquad &= P_{k+1}\Delta y_{k+1} + O(\|P_{k+1}\Delta y_{k+1}\|^2) = P_{k+1}\Delta y_{k+1} + O(h\|P_{k+1}\Delta y_{k+1}\|).
\end{aligned}$$

Similarly, from (4.11) we have

$$(4.16)\qquad \Delta y_k = P_k\Delta y_k + O(h\|P_k\Delta y_k\|).$$

From (4.9) we have

$$Q_{k+1}\Delta z_{k+1} = O(\|\Delta y_{k+1}\| + h\|\Delta z_{k+1}\|) = O(\|P_{k+1}\Delta y_{k+1}\| + h\|\Delta z_{k+1}\|).$$

Therefore,

$$\begin{aligned}
\Delta z_{k+1} &= P_{k+1}\Delta z_{k+1} + Q_k\Delta z_{k+1} = P_{k+1}\Delta z_{k+1} + O(\|\Delta y_{k+1}\| + h\|\Delta z_{k+1}\|)\\
(4.17)\qquad &= P_{k+1}\Delta z_{k+1} + O(\|P_{k+1}\Delta y_{k+1}\| + h\|P_{k+1}\Delta z_{k+1}\|).
\end{aligned}$$

Similarly, from $0 = g_t(t_k,\widehat{y}_k) + g_y(t_k,\widehat{y}_k)\widehat{z}_k$ and $0 = g_t(t_k,\widetilde{y}_k) + g_y(t_k,\widetilde{y}_k)\widetilde{z}_k$ we have

$$(4.18)\qquad \Delta z_k = P_k\Delta z_k + O(\|P_k\Delta y_k\| + h\|P_k\Delta z_k\|).$$

Taking into account the above estimates (4.15)-(4.16)-(4.17)-(4.18) into (4.14) finally leads to the desired result. ☐

Global convergence of the HHT-α method (3.4) can now be proved:

THEOREM 4.5. *Consider the overdetermined system of DAEs (3.1) with initial conditions (y_0, z_0) at t_0 and a_0 satisfying*

$$g(t_0,y_0) = 0,\quad g_t(t_0,y_0) + g_y(t_0,y_0)z_0 = 0,\quad a_0 - a(t_0+\alpha h) = O(h).$$

Then the HHT-α solution (y_n, z_n, a_n) to the system of equations (3.4) satisfies for $0 \leq h \leq h_0$ and $t_n - t_0 = nh \leq Const$

$$y_n - y(t_n) = O(h^2), \quad z_n - z(t_n) = O(h^2), \quad a_n - a(t_n + \alpha h) = O(h^2)\ (n \geq 1),$$

where $(y(t), z(t))$ is the exact solution to (3.1) at t passing through (y_0, z_0) at t_0 and $a(t)$ is given by (3.1c).

Proof. We consider two neighboring HHT-α approximations $(y_k^{k_1}, z_k^{k_1}, a_k^{k_1})_{k=k_1}^n$, $(y_k^{k_1-1}, z_k^{k_1-1}, a_k^{k_1-1})_{k=k_1}^n$ with $k_1 = 1, \dots, n$ and we denote their difference by $\Delta y_k := y_k^{k_1} - y_k^{k_1-1}$, $\Delta z_k := z_k^{k_1} - z_k^{k_1-1}$, $\Delta a_k := a_k^{k_1} - a_k^{k_1-1}$. We assume that $\Delta y_k = O(h)$, $\Delta z_k = O(h)$, $\Delta a_k = O(h)$. These assumptions can be justified by induction, see below. For $k = k_1$, $\Delta y_{k_1}, \Delta z_{k_1}, \Delta a_{k_1}$ are just the local error (4.2) of the HHT-α method (3.4) with $(y_{k_1}^{k_1}, z_{k_1}^{k_1})$ being the exact solution passing through $(y_{k_1-1}^{k_1-1}, z_{k_1-1}^{k_1-1})$ and $(y_{k_1}^{k_1-1}, z_{k_1}^{k_1-1}, a_{k_1}^{k_1-1})$ being the HHT-α numerical approximation from the same point. The HHT-α approximations satisfy the constraints (3.1d)-(3.1e) and we get by application of Theorem 4.4 for $k = k_1, \dots, n-1$

$$\begin{aligned}
P_{k+1}\Delta y_{k+1} &= P_k\Delta y_k + hP_k\Delta z_k + (1/2 - \beta)\, h^2 P_k \Delta a_k \\
&\quad + O(h\|P_k\Delta y_k\| + h^2\|P_k\Delta z_k\| + h^3\|P_k\Delta a_k\| + h^3\|Q_k\Delta a_k\|), \\
P_{k+1}\Delta z_{k+1} &= P_k\Delta z_k + (1-\gamma)\, hP_k\Delta a_k \\
&\quad + O(h\|P_k\Delta y_k\| + h\|P_k\Delta z_k\| + h^2\|P_k\Delta a_k\| + h^2\|Q_k\Delta a_k\|), \\
hP_{k+1}\Delta a_{k+1} &= O(h\|P_k\Delta y_k\| + h\|P_k\Delta z_k\| + h^2\|P_k\Delta a_k\| + h^2\|Q_k\Delta a_k\|), \\
hQ_{k+1}\Delta a_{k+1} &= O(h\|P_k\Delta y_k\| + h\|P_k\Delta z_k\| + h^2\|P_k\Delta a_k\| + h^2\|Q_k\Delta a_k\|).
\end{aligned}$$

Taking a norm of these expressions leads to the estimates

$$\begin{bmatrix} \|P_{k+1}\Delta y_{k+1}\| \\ \|P_{k+1}\Delta z_{k+1}\| \\ h\|P_{k+1}\Delta a_{k+1}\| \\ h\|Q_{k+1}\Delta a_{k+1}\| \end{bmatrix} \leq M \begin{bmatrix} \|P_k\Delta y_k\| \\ \|P_k\Delta z_k\| \\ h\|P_k\Delta a_k\| \\ h\|Q_k\Delta a_k\| \end{bmatrix}$$

with matrix

$$M := \begin{bmatrix} 1+O(h) & h+O(h^2) & h|\frac{1}{2}-\beta| + O(h^2) & O(h^2) \\ O(h) & 1+O(h) & |1-\gamma| + O(h) & O(h) \\ O(h) & O(h) & O(h) & O(h) \\ O(h) & O(h) & O(h) & O(h) \end{bmatrix}.$$

Defining the matrix

$$T := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -|1-\gamma| & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

we can first transform the matrix M by a similarity transformation to the form

$$N := T^{-1}MT = \begin{bmatrix} 1+O(h) & h+O(h^2) & h(|\frac{1}{2}-\beta| - |1-\gamma|) + O(h^2) & O(h^2) \\ O(h) & 1+O(h) & O(h) & O(h) \\ O(h) & O(h) & O(h) & O(h) \\ O(h) & O(h) & O(h) & O(h) \end{bmatrix}.$$

ETNA
Kent State University
etna@mcs.kent.edu

For the matrix N there is a first linear invariant subspace $V_{12} \subset \mathbb{R}^4$ associated to the two eigenvalues $\mu_1 = 1 + O(h)$ and $\mu_2 = 1 + O(h)$. This subspace V_{12} is of the form $V_{12} = \text{span}(e_1 + O(h), e_2 + O(h))$ where $e_1 := (1,0,0,0)^T \in \mathbb{R}^4$ and $e_2 := (0,1,0,0)^T \in \mathbb{R}^4$. For the matrix N there is also a second linear invariant subspace $V_{34} \subset \mathbb{R}^4$ associated to the two eigenvalues $\mu_3 = 0 + O(h)$ and $\mu_4 = 0 + O(h)$. This subspace V_{34} is of the form $V_{34} = \text{span}(e_3 + O(h), e_4 + O(h))$ where $e_3 := (0,0,1,0)^T \in \mathbb{R}^4$ and $e_4 := (0,0,0,1)^T \in \mathbb{R}^4$. Therefore, there is a transformation $V = I + O(h)$ with inverse $V^{-1} = I + O(h)$ such that $NV = VB$ with B block-diagonal, i.e.,

$$B := V^{-1}NV = \begin{bmatrix} 1+O(h) & O(h) & 0 & 0 \\ O(h) & 1+O(h) & 0 & 0 \\ 0 & 0 & O(h) & O(h) \\ 0 & 0 & O(h) & O(h) \end{bmatrix}.$$

For $2 \le m \le n$, from $mh \le nh \le Const$ we obtain

$$M^m = TVB^mV^{-1}T^{-1} = \begin{bmatrix} O(1) & O(1) & O(1) & O(h) \\ O(1) & O(1) & O(1) & O(h) \\ O(h) & O(h) & O(h) & O(h^2) \\ O(h) & O(h) & O(h) & O(h^2) \end{bmatrix},$$

giving

$$\begin{bmatrix} \|P_n\Delta y_n\| \\ \|P_n\Delta z_n\| \\ h\|P_n\Delta a_n\| \\ h\|Q_n\Delta a_n\| \end{bmatrix} \le \begin{bmatrix} O(1) & O(1) & O(1) & O(h) \\ O(1) & O(1) & O(1) & O(h) \\ O(h) & O(h) & O(h) & O(h^2) \\ O(h) & O(h) & O(h) & O(h^2) \end{bmatrix} \begin{bmatrix} \|P_{n-m}\Delta y_{n-m}\| \\ \|P_{n-m}\Delta z_{n-m}\| \\ h\|P_{n-m}\Delta a_{n-m}\| \\ h\|Q_{n-m}\Delta a_{n-m}\| \end{bmatrix}.$$

By Theorem 4.4 we have

$$Q_k\Delta y_k = O(\|P_k\Delta y_k\|), \qquad Q_k\Delta z_k = O(\|P_k\Delta y_k\| + \|P_k\Delta z_k\|).$$

Hence since $\Delta y_k = P_k\Delta y_k + Q_k\Delta y_k$ and $\Delta z_k = P_k\Delta z_k + Q_k\Delta z_k$, we get

$$\begin{bmatrix} \|\Delta y_n\| \\ \|\Delta z_n\| \\ \|\Delta a_n\| \end{bmatrix} \le C \begin{bmatrix} \|\Delta y_{n-m}\| + \|\Delta z_{n-m}\| + h\|\Delta a_{n-m}\| \\ \|\Delta y_{n-m}\| + \|\Delta z_{n-m}\| + h\|\Delta a_{n-m}\| \\ \|\Delta y_{n-m}\| + \|\Delta z_{n-m}\| + h\|\Delta a_{n-m}\| \end{bmatrix}. \tag{4.19}$$

First, we consider the HHT-α solution using the exact value $a(t_0 + \alpha h)$. We denote it by $(\overline{y}_k, \overline{z}_k, \overline{a}_k)_{k=0}^n$. For $k = 0$ we have $\overline{y}_0 = y_0$, $\overline{z}_0 = z_0$, and $\overline{a}_0 = a(t_0 + \alpha h)$. Taking $m := k_1$ in (4.19) leads to

$$\|\Delta y_n\| \le c_y h^3, \qquad \|\Delta z_n\| \le c_z h^3, \qquad \|\Delta a_n\| \le c_a h^3.$$

The assumptions $\Delta y_k = O(h)$, $\Delta z_k = O(h)$, $\Delta a_k = O(h)$ are thus justified by induction on k. Summing up these estimates we obtain

$$\|y(t_n) - \overline{y}_n\| \le \sum_{k_1=1}^{n} \|y_n^{k_1} - y_n^{k_1-1}\| \le c_y n h^3 \le C_y h^2,$$

$$\|z(t_n) - \overline{z}_n\| \le \sum_{k_1=1}^{n} \|z_n^{k_1} - z_n^{k_1-1}\| \le c_z n h^3 \le C_z h^2,$$

$$\|a(t_n + \alpha h) - \overline{a}_n\| \le \sum_{k_1=1}^{n} \|a_n^{k_1} - a_n^{k_1-1}\| \le c_a n h^3 \le C_a h^2.$$

ETNA
Kent State University
etna@mcs.kent.edu

Now, suppose that a_0 satisfies $a_0 = a(t_0 + \alpha h) + O(h)$. We denote the corresponding HHT-α solution using this approximate value of a_0 by $(y_k, z_k, a_k)_{k=0}^{n}$. We want to estimate $\|y_n - \overline{y}_n\|$, $\|z_n - \overline{z}_n\|$, and $\|a_n - \overline{a}_n\|$. Using (4.19) for $m = n$, since $\overline{y}_0 = y_0$ and $\overline{z}_0 = z_0$, we simply obtain

$$\begin{bmatrix} \|y_n - \overline{y}_n\| \\ \|z_n - \overline{z}_n\| \\ \|a_n - \overline{a}_n\| \end{bmatrix} \leq C \begin{bmatrix} h\|a_0 - \overline{a}_0\| \\ h\|a_0 - \overline{a}_0\| \\ h\|a_0 - \overline{a}_0\| \end{bmatrix} = O(h^2)$$

since $a_0 - \overline{a}_0 = a_0 - a(t_0 + \alpha h) = O(h)$. The assumptions $y_k - \overline{y}_k = O(h)$, $z_k - \overline{z}_k = O(h)$, $a_k - \overline{a}_k = O(h)$ are also justified by induction on k. By combining the above estimates we finally we get the desired result

$$\begin{aligned} \|y_n - y(t_n)\| &= \|y_n - \overline{y}_n\| + \|\overline{y}_n - y(t_n)\| = O(h^2), \\ \|z_n - z(t_n)\| &= \|z_n - \overline{z}_n\| + \|\overline{z}_n - z(t_n)\| = O(h^2), \\ \|a_n - a(t_n + \alpha h)\| &= \|a_n - \overline{a}_n\| + \|\overline{a}_n - a(t_n + \alpha h)\| = O(h^2). \end{aligned}$$

Remark that the proof of this Theorem remains valid with variable stepsize provided the values a_k are corrected for example by (2.3). $\square$

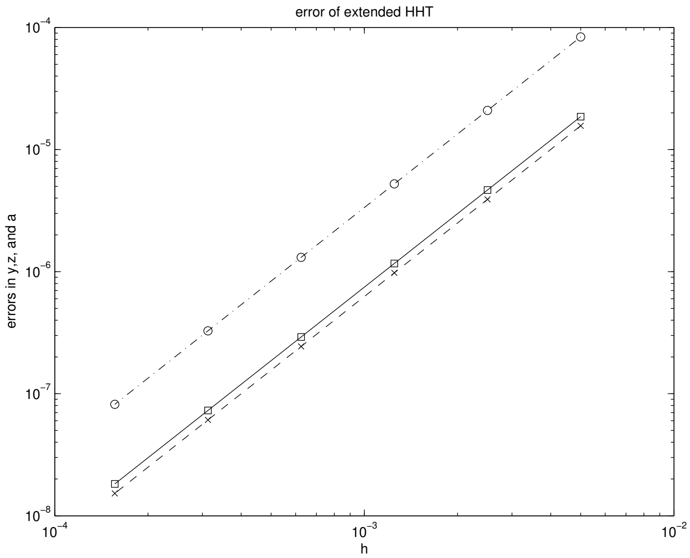

FIG. 5.1. *Global errors* $\|y_n - y(t_n)\|_2$ ($\square$), $\|z_n - z(t_n)\|_2$ ($\circ$), $\|a_n - a(t_n + \alpha h)\|_2$ ($\times$) *for first test problem at* $t_n = 1$ *and the extended HHT-*α *method* $(\alpha = -0.15, b = 0.3)$. *One observes global convergence of order* 2 *in* h.

5. Numerical experiments.

5.1. A first test problem. We consider the following mathematical test problem

$$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \qquad \begin{bmatrix} z_1' \\ z_2' \end{bmatrix} = \begin{bmatrix} y_1 z_2 + 2y_2 z_1 + e^t y_1 \lambda_1 \\ \frac{1}{2} y_2 z_2 - 2y_1 z_1 y_2 z_2 + y_2 \lambda_1^2 \end{bmatrix},$$
$$\begin{bmatrix} 0 \end{bmatrix} = \begin{bmatrix} y_1^2 y_2 - 1 \end{bmatrix}, \qquad \begin{bmatrix} 0 \end{bmatrix} = \begin{bmatrix} 2y_1 y_2 z_1 + y_1^2 z_2 \end{bmatrix}.$$

ETNA
Kent State University
etna@mcs.kent.edu

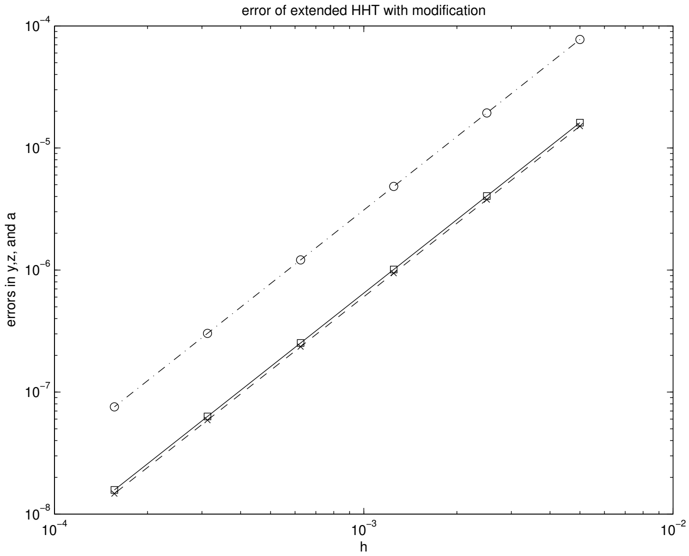

FIG. 5.2. *Global errors* $\|y_n - y(t_n)\|_2$ ($\square$), $\|z_n - z(t_n)\|_2$ ($\circ$), $\|a_n - a(t_n + \alpha h)\|_2$ ($\times$) *for first test problem at* $t_n = 1$ *and the extended HHT-*α *method* $(\alpha = -0.15, b = 0.3)$ *with modification (3.5). One observes global convergence of order* 2 *in* h.

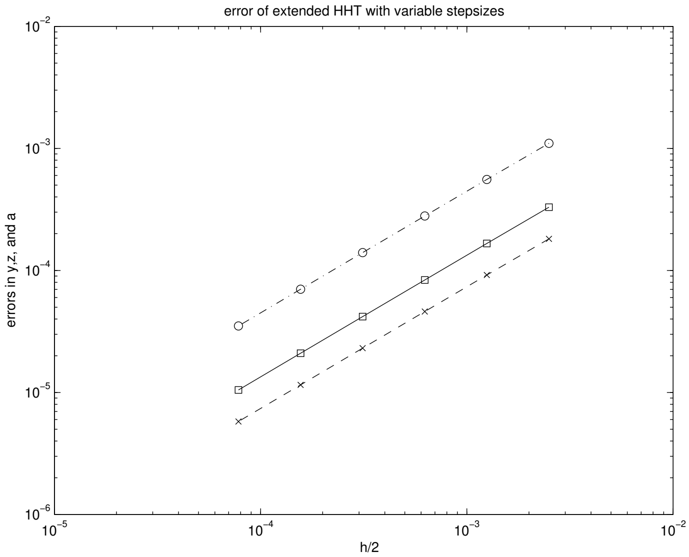

FIG. 5.3. *Global errors* $\|y_n - y(t_n)\|_2$ ($\square$), $\|z_n - z(t_n)\|_2$ ($\circ$), $\|a_n(h_{n-1}) - a(t_n + \alpha h_{n-1})\|_2$ ($\times$) *for first test problem at* $t_n = 1$ *and the extended HHT-*α *method* $(\alpha = -0.15, b = 0.3)$ *with variable stepsize and unmodified* a_n. *One observes global convergence of order* 1 *in* h.

Notice that these equations are nonlinear in λ_1. Consistent initial conditions at $t_0 = 0$ are given by

$$\begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} z_1(0) \\ z_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \qquad \begin{bmatrix} \lambda_1(0) \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}.$$

ETNA
Kent State University
etna@mcs.kent.edu

The exact solution to this system of DAEs is given by

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} e^t \\ e^{-2t} \end{bmatrix}, \qquad \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} e^t \\ -2e^{-2t} \end{bmatrix}, \qquad \begin{bmatrix} \lambda_1(t) \end{bmatrix} = \begin{bmatrix} e^{-t} \end{bmatrix}.$$

We have applied the extended HHT-α method (3.4) with parameters $\alpha = -0.15$ and $b = 0.3$ for various stepsize h. We observe global convergence of order 2 at $t = 1$ in Fig. 5.1 as expected from Theorem 4.5. In Fig. 5.2 we have repeated the same numerical experiment by simply replacing y_1 in R_1 with (3.5). We still observe global convergence of order 2. In Fig. 5.3 we have applied the HHT-α method with variable stepsize alternating between $h/3$ and $2h/3$. We have plotted in Fig. 5.3 the error versus the average stepsize $h/2$. We observe a reduction to convergence of order one as expected from the remarks in Section 2. To reestablish second order convergence for variable stepsize we have made use of the modification (2.3) for a_n, i.e.,

$$a_n := f(t_n, y_n, z_n) + \frac{h_n}{h_{n-1}}(a_n(h_{n-1}) - f(t_n, y_n, z_n)).$$

We have applied again the HHT-α method with variable stepsize alternating between $h/3$ and $2h/3$ using this modification. This time we observe second order global convergence in Fig. 5.4.

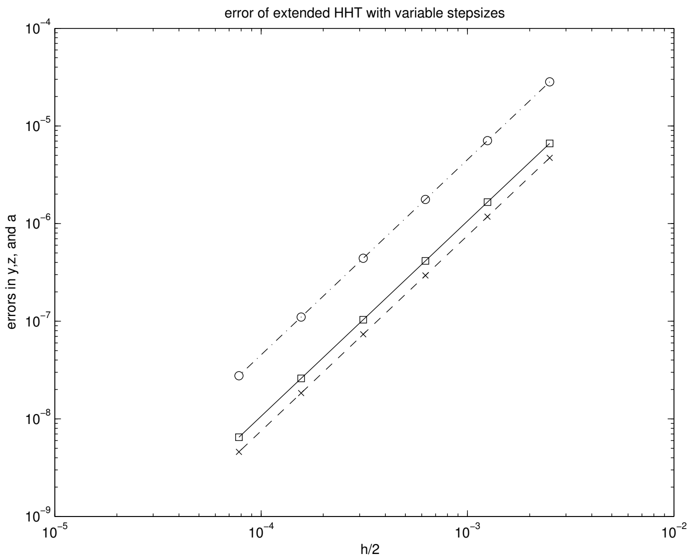

FIG. 5.4. *Global errors* $\|y_n - y(t_n)\|_2$ ($\square$), $\|z_n - z(t_n)\|_2$ ($\circ$), $\|a_n(h_{n-1}) - a(t_n + \alpha h_{n-1})\|_2$ ($\times$) *for first test problem at* $t_n = 1$ *and the extended HHT-*α *method* ($\alpha = -0.15, b = 0.3$) *with variable stepsize and modified* a_n *(2.3). One observes global convergence of order* 2 *in* h.

5.2. A pendulum model. The pendulum model in Fig. 5.5 was used to carry out a second set of numerical experiments. Using the notation $z_i = y_i'$ $(i = 1, 2, 3)$, the constrained equations of motion associated with this model are

$$\begin{bmatrix} mz_1' \\ mz_2' \\ \frac{mL^2}{3} z_3' \end{bmatrix} = \begin{bmatrix} 0 \\ -mg \\ -cz_3 - k \cdot \left(y_3 - \frac{3\pi}{2}\right) \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ L\sin(y_3) & -L\cos(y_3) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix},$$

ETNA
Kent State University
etna@mcs.kent.edu

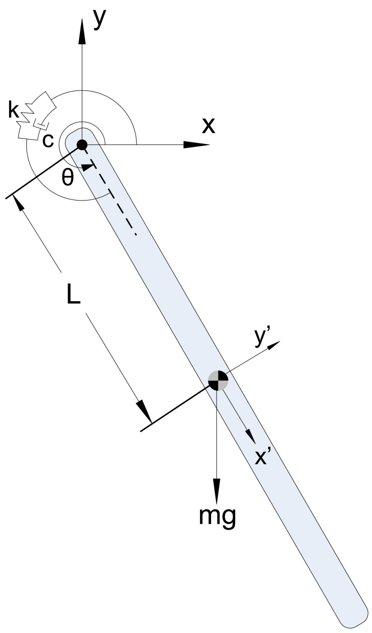

FIG. 5.5. *Pendulum model.*

while the constraint equations at the position and velocity levels are

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} y_1 - L\cos(y_3) \\ y_2 - L\sin(y_3) \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} z_1 + L\sin(y_3)z_3 \\ z_2 - L\cos(y_3)z_3 \end{bmatrix}.$$

The parameters associated with this model are as follows: mass $m = 5$, length $L = 2$, spring stiffness $k = 3000$, damping coefficient $c = 100$, gravitational acceleration $g = 9.81$. All units used herein are SI units. The evolution of the pendulum angle $y_3 = \theta$ on the time interval $[0, 4]$ is shown in Fig. 5.6.

In the numerical experiments, the pendulum is started from an initial position that corresponds to $y_3 = 3\pi/2$, and $z_3 = 10$. A reference solution was generated by applying an explicit Runge-Kutta method of order 4 (RK4) with a small constant stepsize $h = 0.00001$. The explicit integrator RK4 was used in conjunction with an equivalent underlying ODE problem that provided directly the time evolution of y_3

$$y_3' = z_3, \qquad \frac{4mL^2}{3}z_3' + cz_3 + k\cdot\left(y_3 - \frac{3\pi}{2}\right) + mgL\cos(y_3) = 0.$$

Figs. 5.7 and 5.8 support the convergence results obtained in Theorem 4.5. The global error in y_3 and z_3 at time $t = 2$ is plotted in these figures versus a series of stepsize used for integration. The plots confirm that the global errors $|y_{3,n} - y_3(t_n)|$ and $|z_{3,n} - z_3(t_n)|$ associated with the extended HHT-α method (3.4) are of order two. Note that Figs. 5.7 and 5.8 report results for $\alpha = 0, b = 0$ and $\alpha = -0.3, b = 0$ respectively.

6. Extension of the HHT-α method to DAEs with index 2 constraints. Consider semi-explicit index 2 DAEs of the form

$$y' = z, \tag{6.1a}$$

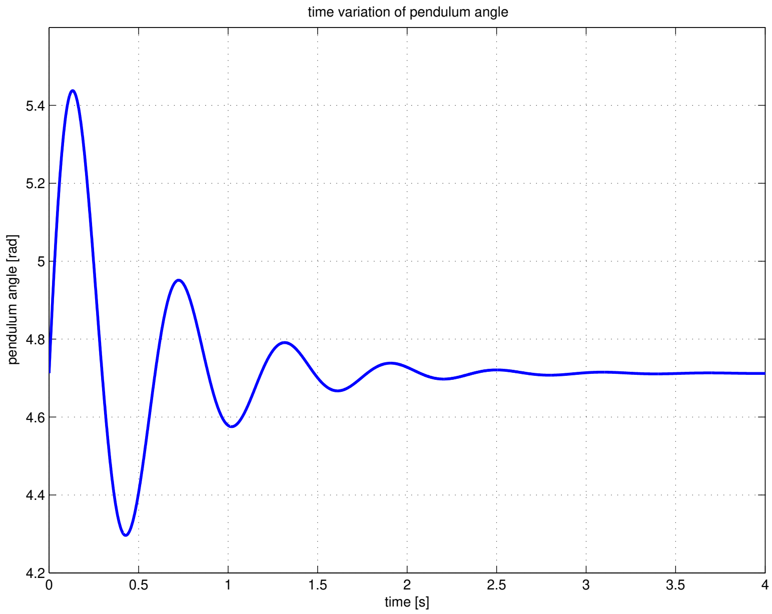

FIG. 5.6. *Time evolution of pendulum angle* y_3.

$$z' = a + r(t, y, z, \psi), \tag{6.1b}$$
$$a = f(t, y, z), \tag{6.1c}$$
$$0 = k(t, y, z), \tag{6.1d}$$

where we assume $k_z(t, y, z)r_\psi(t, y, z, \psi)$ is invertible in the region of interest. We can consider for example $k(t, y, z) = g_t(t, y) + g_y(t, y)z$ from (3.1e) and $r(t, y, z, \psi) = r(t, y, \psi)$ from (3.1b). We propose a generalization of HHT-α methods for the system (6.1) similar to (3.4),

$$\begin{aligned} y_1 &= y_0 + hz_0 + \frac{h^2}{2}((1-2\beta)a_0 + 2\beta a_1) + \frac{h^2}{2}R_{1/2}, \\ z_1 &= z_0 + h((1-\gamma)a_0 + \gamma a_1) + hR_{1/2}, \\ a_1 &= (1+\alpha)f(t_1, y_1, z_1) - \alpha f(t_0, y_0, z_0), \end{aligned}$$

where

$$R_{1/2} = r\left(t_0 + \frac{h}{2}, y_0 + \frac{h}{2}z_0, \frac{1}{2}(z_0 + z_1), \Psi_{1/2}\right).$$

The algebraic variable $\Psi_{1/2}$ is determined by the constraint

$$0 = k(t_1, y_1, z_1).$$

Hence, the numerical solution satisfies the constraint (6.1d) at each timestep. By replacing the expression for z_1 explicitly in this equation, we obtain equivalently

$$0 = \frac{1}{h}k(t_1, y_1, z_0 + h\left((1-\gamma)a_0 + \gamma a_1\right) + hR_{1/2}).$$

ETNA
Kent State University
etna@mcs.kent.edu

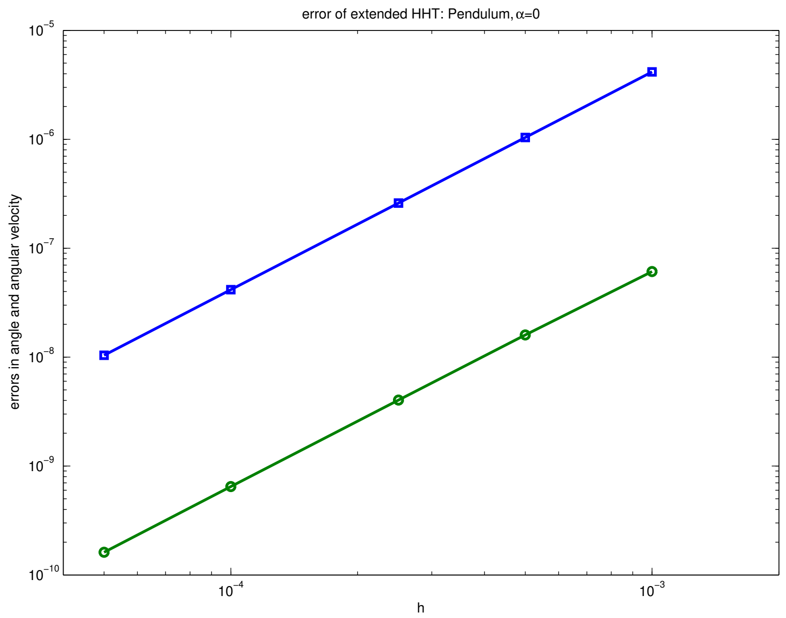

FIG. 5.7. *Global errors* $|y_{3,n} - y_3(t_n)|$ ($\square$), $|z_{3,n} - z_3(t_n)|$ ($\circ$) *at* $t_n = 2$ *for a simple pendulum and the extended HHT-*α *method* $(\alpha = 0, b = 0)$. *One observes global convergence of order* 2 *in* h.

Existence and uniqueness of the numerical solution $(y_1, z_1, a_1, \Psi_{1/2})$ is ensured and can be shown by application of the implicit function theorem. When $k(t, y, z)$ is linear in z and $r(t, y, z, \psi)$ is linear in ψ we obtain a linear equation for $\Psi_{1/2}$. However, since generally $k(t, y, z)$ or $f(t, y, z)$ are nonlinear in y, we generally have a system of nonlinear equations to solve in terms of y_1. If $k(t, y, z)$ and $f(t, y, z)$ are linear in y and z, and if $r(t, y, z, \psi)$ is linear in ψ, we obtain a system of linear equations for $(y_1, z_1, a_1, \Psi_{1/2})$. Global second order convergence of the new extended HHT-α method can be proved in a similar way as for (3.1).

7. Conclusion. In this paper we have presented second order extensions of the HHT-α method for systems of ODAEs with index 3 and index 2 constraints arising for example in mechanics. We have given detailed mathematical proofs of convergence of extensions of the HHT-α method for semi-explicit ODAEs with index 3 constraints and underlying index 2 constraints. We have taken into account the structure of the equations to extend the HHT-α method to DAEs in order to keep its second order accuracy. We have also proposed an elementary way to preserve the second order of the HHT-α method when using variable stepsize, a technique which is also relevant for the HHT-α method when applied to ODEs. The HHT-α method and its extensions to DAEs is relatively simple to express and to implement. However, its analysis in the context of DAEs was found to be surprisingly difficult.

After the submission of this manuscript, we learned about a similar extension of the generalized-α method found independently by Lunk and Simeon [10]. They consider problems of the form (3.1) with $r(t, y, \lambda)$ linear in λ, whereas in our paper $r(t, y, \lambda)$ may be nonlinear. Their extension is slightly different, when $r(t, y, \lambda) = r(t, y)\lambda$ they replace in (3.4a) the term $(1 - b)R_0 + bR_1 = (1 - b)r(t_0, y_0)\Lambda_0 + br(t_1, y_1)\Lambda_1$ by $((1 - b)r(t_0, y_0) + br(t_1, y_1))\Lambda_0$.

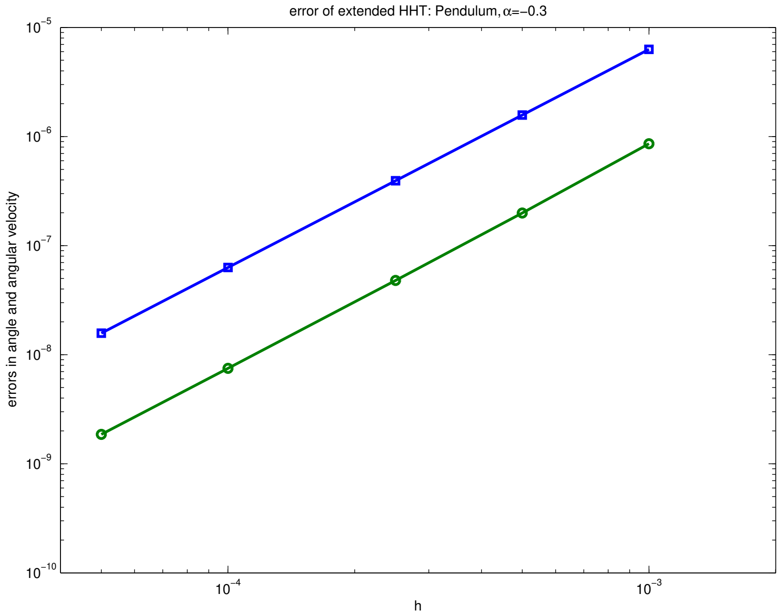

FIG. 5.8. *Global errors* $|y_{3,n} - y_3(t_n)|$ $(\square)$*,* $|z_{3,n} - z_3(t_n)|$ $(\circ)$ *at* $t_n = 2$ *for a simple pendulum and the extended HHT-*α *method* $(\alpha = -0.3, b = 0)$ *One observes global convergence of order* 2 *in* h*.*

REFERENCES

[1] K. E. BRENAN, S. L. CAMPBELL, AND L. R. PETZOLD, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, Second ed., SIAM Classics in Appl. Math., SIAM, Philadelphia, 1996.

[2] A. CARDONA AND M. GÉRADIN, *Time integration of the equations of motion in mechanism analysis*, Comput. & Structures, 33 (1989), pp. 801–820.

[3] J. CHUNG AND G. M. HULBERT, *A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-*α *method*, J. Appl. Mech., 60 (1993), pp. 371–375.

[4] S. ERLICHER, L. BONAVENTURA, AND O. S. BURSI, *The analysis of the generalized-*α *method for nonlinear dynamics problems*, Comput. Mech., 28 (2002), pp. 83–104.

[5] C. W. GEAR, G. K. GUPTA, AND B. J. LEIMKUHLER, *Automatic integration of the Euler-Lagrange equations with constraints*, J. Comput. Appl. Math., 12–13 (1985), pp. 77–90.

[6] H. M. HILBER, T. J. R. HUGHES, AND R. L. TAYLOR, *Improved numerical dissipation for time integration algorithms in structural dynamics*, Earthquake Engng Struct. Dyn., 5 (1977), pp. 283–292.

[7] T. J. R. HUGHES, *Finite Element Method - Linear Static and Dynamic Finite Element Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

[8] L. O. JAY, *Symplectic partitioned Runge-Kutta methods for constrained Hamiltonian systems*, SIAM J. Numer. Anal., 33 (1996), pp. 368–387.

[9] ———, *Structure preservation for constrained dynamics with super partitioned additive Runge-Kutta methods*, SIAM J. Sci. Comput., 20 (1998), pp. 416–446.

[10] C. LUNK AND B. SIMEON, *Solving constrained mechanical systems by the family of Newmark and* α*-methods*, ZAMM, 86 (2006), pp. 772–784.

[11] D. NEGRUT, R. RAMPALLI, G. OTTARSSON, AND A. SAJDAK, *On an implementation of the HHT method in the context of index 3 differential algebraic equations of multibody dynamics*, ASME J. Comp. Nonlin. Dyn., 2 (2007), pp. 73–85.

[12] J. YEN, L. PETZOLD, AND S. RAHA, *A time integration algorithm for flexible mechanism dynamics: The DAE-alpha method*, Comput. Methods Appl. Mech. Engrg., 158 (1998), pp. 341–355.