Numerische Mathematik manuscript No.
(will be inserted by the editor)

Reversible methods of Runge-Kutta type for Index-2 DAEs

R.P.K. Chan[1], **P. Chartier**[2], **A. Murua**[3]

[1] Department of Mathematics, Tamaki Campus, The University of Auckland, Private Bag 92019, Auckland, NEW ZEALAND.
[2] INRIA Rennes, Campus de Beaulieu, 35042 Rennes Cedex, FRANCE.
[3] Konputazio Zientziak eta A.A. saila, Informatika Fakultatea, EHU, Donostia, SPAIN.

December 17, 2003

Summary A new interpretation of Runge-Kutta methods for differential algebraic equations (DAEs) of index 2 is presented, where a step of the method is described in terms of a smooth map (smooth also with respect to the stepsize). This leads to a better understanding of the convergence behavior of Runge-Kutta methods that are not stiffly accurate. In particular, our new framework allows for the unified study of two order-improving techniques for symmetric Runge-Kutta methods (namely post-projection and symmetric projection) specially suited for solving reversible index-2 DAEs.

AMS Classification : 65L05, 65L06

Keywords : differential algebraic equations, index 2, reversibility, Runge-Kutta methods, symmetric projection, backward analysis.

1 Introduction

The main focus of this paper is the integration of *reversible index-2 differential-algebraic equations* with Runge-Kutta methods. More specifically, we consider index-2 DAEs in semi-explicit form[1]

$$\begin{cases} y' = f(y,z), & f : \mathbb{R}^{M+N} \to \mathbb{R}^{M}, \\ 0 \; = g(y), & g : \mathbb{R}^{M} \quad\;\; \to \mathbb{R}^{N}, \end{cases} \tag{1}$$

[1] See [HLR89] for a definition of the notion of index.

where, as is usual, $g_y(y)f_z(y,z)$ is assumed to be invertible in a neighborhood of the exact solution, and the initial values y_0, z_0 are assumed to be consistent, that is,

$$g(y_0) = 0, \qquad g_y(y_0)f(y_0, z_0) = 0.$$

We suppose in addition that there exist isomorphisms $\rho : \mathbb{R}^M \to \mathbb{R}^M$, $\rho^2 = I_M$ and $\widetilde{\rho} : \mathbb{R}^N \to \mathbb{R}^N$ such that

$$\begin{cases} \rho f(y,z) = -f(\rho y, z), \\ g(\rho y) \quad = \widetilde{\rho} g(y). \end{cases} \tag{2}$$

Note that the special case, frequently arising in practice, of mechanical systems with holonomic constraints and an Hamiltonian function of the form $H(q,p) = \frac{1}{2}p^T M(q)^{-1}p + U(q)$, can be put into the form (1) and satisfy (2). The equations of motion of nonholonomic mechanics can often be put into this form as well.

The definition we give here of reversible index-2 DAEs is related to reversible ODEs as follows : under some minor additional assumptions, equations (1) can be rewritten as

$$y' = f(y, \theta(y)), \tag{3}$$
$$z = \theta(y), \tag{4}$$

where θ is the smooth function implicitely defined as the solution of $g_y(y)f(y,\theta(y)) = 0$. It is then straightforward to check that, if condition (2) holds, $f(\rho y, \theta(\rho y)) = -\rho f(y, \theta(y))$. In that case, (3) is a ρ-reversible ODE system, that is,

$$\rho \circ \varphi_t = \varphi_{-t} \circ \rho,$$

where φ_t denotes the exact y-flow of the ODE system (3).

It is known [Sto88] that the numerical flow defined by the application of a symmetric Runge-Kutta method to a ρ-reversible ODE system is ρ-reversible. Backward error analysis can be applied in that case to show that [HLW02] the numerical flow is exponentially close to the exact flow of a ρ-reversible ODE system (which is a perturbation of the original one).

This suggests that symmetric Runge-Kutta schemes might be of interest for the numerical integration of reversible index-2 DAEs and amongst them the important class of Gauss collocation methods. Although such methods suffer from an important order reduction [HW96], it has been shown in [CCM02b] to be of local nature, so that once the Runge-Kutta solutions have been obtained, the original order of convergence can be restored by performing a suitable projection (*post-projection*) onto the manifold $\mathcal{M} = \{y \in \mathbb{R}^M; g(y) = 0\}$.

Another possibility is to adapt to index-2 DAEs the idea of *symmetric projection* proposed in [Hai00, Hai01] for ODEs.

In a unified framework, we thus present a new interpretation of Runge-Kutta schemes with invertible matrix, where one step of the method is interpreted in terms of a smooth map (smooth also in the stepsize). In addition to shedding new light toward the understanding of the convergence behavior of general Runge-Kutta methods, by establishing a connection with the concept of effective order for ODEs [But69], it immediately allows the application of standard backward error analysis for ODEs. Moreover, this approach greatly simplifies the analysis of symmetrically projected Runge-Kutta methods for index-2 DAEs.

In Section 2, we shall first present an overview of the main tools used in the paper and the motivations behind them. Section 3 will introduce our new framework for the study of the application of Runge-Kutta methods to index-2 DAEs. The convergence analysis for the special case of collocation methods will be given in Section 4, while a backward analysis will be sketched in Section 5 whose results are particularly useful in understanding the application of Runge-Kutta methods (in particular with post-projection) to reversible index-2 DAEs. In Section 6, we will study the special case of *symmetrically projected methods*. In Section 7, we will report on some numerical experiments aimed at illustrating the convergence results obtained here as well as the properties of symmetric methods, with or without symmetric projection, for long-term integration.

2 Setting up the problem : main notation and motivations

As usual when solving index-2 DAEs numerically, we will assume that $g_y(y)f_z(y,z)$ is invertible in a neighborhood of the manifold $\mathcal{M}$. More precisely, we will suppose that there exists a bounded and connected open neighborhood $\mathcal{V}$ of $\mathcal{M}$ where we can define a smooth map θ, from $\mathcal{V}$ to $\mathbb{R}^N$, such that

$$\forall\, y \in \mathcal{V}, \exists!\, z = \theta(y) \in \mathbb{R}^N, g_y(y)f(y,z) = 0 \text{ and } \det[g_y(y)f_z(y,z)] \neq 0.$$

Eventually, we will refer to a Runge-Kutta scheme with coefficient matrix A and weight vector b by $\mathcal{R} = (A, b)$. We hereafter assume that A is invertible. The numerical flow of (1) corresponding to the Runge-Kutta scheme $\mathcal{R} = (A, b)$ is the map $y^* = \Phi_h(y)$ defined by

$$y^* = y + h(b^T \otimes I_M)F(Y,Z), \tag{5}$$

where the internal stage values Y, Z are implicitly defined by

$$Y = (e \otimes y) + h(A \otimes I_M)F(Y,Z), \tag{6}$$
$$0 = G(Y), \tag{7}$$

where $F(Y,Z)$ and $G(Y)$ are the vector-functions with components $f(Y_i,Z_i)$ and $g(Y_i)$, $i = 1,\dots,s$. Here the question of whether Φ_h is well defined for sufficiently small h arises. It is clear that, for $h = 0$, (6)-(7) has no solution if $y \notin \mathcal{M}$. This causes difficulties[2] in studying the successive application of Φ_h as $h \to 0$, especially when the method is not *stiffly accurate*[3], since the numerical solution

$$y_n = \Phi_h(y_{n-1}), \quad n \geq 1 \tag{8}$$

will not in general lie on the constraint manifold $\mathcal{M}$ even if $y_0 \in \mathcal{M}$. We will hereafter address this problem in a rigorous setting. We first give an overview.

We consider the standard projection[4] $y = \mathcal{P}_{\mathcal{M}}(\widehat{y})$ onto the manifold $\mathcal{M}$ implicitly defined as follows

$$y = \widehat{y} + f_z(y,\theta(y))\mu, \quad 0 = g(y),$$

and we denote $\mu = \mathcal{P}_{\mathbb{R}^N}(\widehat{y})$. The pair $(y,\mu) \in \mathcal{M} \times \mathbb{R}^N$ is uniquely defined for each $\widehat{y} \in \mathbb{R}^M$ sufficiently close to $\mathcal{M}$. This can be viewed as a change of variables from $\mathbb{R}^M$ to $\mathcal{M} \times \mathbb{R}^N$.

The key to our approach is to rewrite the numerical flow Φ_h in the new variables (y,λ). That is, denoting $B_N(0,\alpha) \subset \mathbb{R}^N$ the open ball of center 0 and radius α (where α is an arbitrarily fixed positive value), we define a new map Ψ_h as follows

Definition 1 *The map Ψ_h parametrised by $h \in \mathbb{R}$ is the function*

$$\begin{aligned}\Psi_h \ :\ & \mathcal{M} \times B_N(0,\alpha) \to \mathcal{M} \times \mathbb{R}^N \\ & (y,\lambda) \mapsto (y^*,\lambda^*) = (\Psi_h^{[1]}(y,\lambda), \Psi_h^{[2]}(y,\lambda)),\end{aligned}$$

defined by the equations

$$\widehat{y} = y - h^2 f_z(y,\theta(y))\lambda, \tag{9}$$
$$\widehat{y}^* = \Phi_h(\widehat{y}), \tag{10}$$
$$y^* = \widehat{y}^* + h^2 f_z(y^*,\theta(y^*))\lambda^*, \ \ g(y^*) = 0. \tag{11}$$

[2] It is only known [HW96] that (6)-(7) has a unique solution if $g(y) = \mathcal{O}(h^2)$.
[3] See [HW96] for a discussion of stiffly accurate methods.
[4] As opposed to the *symmetric* projection discussed later on.

For each $(y, \lambda) \in \mathcal{M} \times \mathbb{R}^N$, $\Psi_h(y, \lambda)$ represents a smooth curve on $\mathcal{M} \times \mathbb{R}^N$ defined for sufficiently small $h \neq 0$, which can be smoothly extended to $h = 0$.

Eventually, it is worth mentioning that the two maps Φ_h and Ψ_h are conjugate to each other in the sense that the following diagram commutes :

$$\begin{array}{ccccc} & & \Phi_h & & \\ & \widehat{y} & \longmapsto & \widehat{y}^* & \\ \mathcal{P} & \downarrow & & \downarrow & \mathcal{P} \\ & (y,\lambda) & \longmapsto & (y^*,\lambda^*) & \\ & & \Psi_h & & \end{array} \tag{12}$$

As a result, the application of n steps with constant step-size h of the map Ψ_h, starting from a point $(y, 0)$ with y on the manifold $\mathcal{M}$, may be seen as the application of n steps of the Runge-Kutta method Φ_h with a final projection step. This is exactly what we called post-projection in [CCM02b].

3 Smoothness of Ψ_h

The main result of this section is the following theorem :

Theorem 1 *Consider a Runge-Kutta method $\mathcal{R} = (A, b)$ with invertible matrix A. There exists $\tilde{h} > 0$ such that, for any $|h| < \tilde{h}$, Ψ_h is well-defined and smooth on $\mathcal{M} \times B_N(0, \alpha)$. More precisely, when considered as a function of (y, λ, h), Ψ_h is smooth on $\mathcal{U} := \mathcal{M} \times B_N(0, \alpha) \times (-\tilde{h}, \tilde{h})$. Moreover, we have for $(y^*, \lambda^*) = \Psi_h(y, \lambda)$ the estimates*

$$\frac{\partial y^*}{\partial \lambda} = \mathcal{O}(h^3) \text{ and } \frac{\partial \lambda^*}{\partial \lambda} = R(\infty) I_N + \mathcal{O}(h), \tag{13}$$

where $R(\infty) = 1 - b^T A^{-1} e$.

Proof According to the definitions of Φ_h and Ψ_h, (y^*, λ^*) is given by the equations

$$\widehat{y} = y - h^2 f_z(y, \theta(y))\lambda, \tag{14}$$

$$\widehat{Y} = (e \otimes \widehat{y}) + h(A \otimes I_M) F(\widehat{Y}, \widehat{Z}), \tag{15}$$

$$0 = G(\widehat{Y}), \tag{16}$$

$$\widehat{y}^* = \widehat{y} + h(b^T \otimes I_M) F(\widehat{Y}, \widehat{Z}), \tag{17}$$

$$y^* = \widehat{y}^* + h^2 f_z(y^*, \theta(y^*))\lambda^*, \tag{18}$$

$$0 = g(y^*). \tag{19}$$

Rescaling equations (16) and (19) and using the *Implicit Function Theorem*, it is standard[5] to show that, for $(y,\lambda) \in \mathcal{M} \times B_N(0,\alpha)$ and sufficiently small $|h|$, equations (14-15-16-17-18-19) define (y^*,λ^*) as a smooth function of (y,λ,h) with values in a neighborhood of $(y,0)$. The compactness of $\mathcal{M} \times \overline{B}_N(0,\alpha)$ then allows to get the same result on the whole set $\mathcal{U}$ provided $\tilde{h}$ is small enough.

Now, differentiating (15) and (16) with respect to λ gives successively

$$\frac{\partial \widehat{Y}}{\partial \lambda} = -h^2(e \otimes f_z) + h(A \otimes I_M)\left(\{f_y\}\frac{\partial \widehat{Y}}{\partial \lambda} + \{f_z\}\frac{\partial \widehat{Z}}{\partial \lambda}\right),$$

$$0 = \{g_y\}\frac{\partial \widehat{Y}}{\partial \lambda},$$

where $\{f_y\}$, $\{f_z\}$ and $\{g_y\}$ are block-diagonal matrices with diagonal elements $f_y(\widehat{Y}_i,\widehat{Z}_i)$, $f_z(\widehat{Y}_i,\widehat{Z}_i)$ and $g_y(\widehat{Y}_i)$ respectively and where f_z and later on g_y are evaluted at $(y,\theta(y))$. Inserting the first equation into the second one we obtain

$$0 = \{g_y\}(A \otimes I_M)\{f_y\}\frac{\partial \widehat{Y}}{\partial \lambda} + \{g_y\}(A \otimes I_M)\{f_z\}\frac{\partial \widehat{Z}}{\partial \lambda} - h\{g_y\}(e \otimes f_z),$$

so that we get $\frac{\partial \widehat{Z}}{\partial \lambda} = \mathcal{O}(h)$ and $\frac{\partial \widehat{Y}}{\partial \lambda} = \mathcal{O}(h^2)$. Eventually, previous equation becomes

$$0 = \{g_y\}(A \otimes I_M)\{f_z\}\frac{\partial \widehat{Z}}{\partial \lambda} - h\{g_y\}(e \otimes f_z) + \mathcal{O}(h^2),$$

and we easily get

$$\begin{aligned}\frac{\partial \widehat{Z}}{\partial \lambda} &= h(\{g_y\}(A \otimes I_M)\{f_z\})^{-1}\{g_y\}(e \otimes f_z) + \mathcal{O}(h^2),\\ &= h(A^{-1}e \otimes I_N) + \mathcal{O}(h^2), \qquad (20)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial \widehat{Y}}{\partial \lambda} &= -h^2(e \otimes f_z) + h(A \otimes I_M)(I_s \otimes f_z + \mathcal{O}(h))\frac{\partial \widehat{Z}}{\partial \lambda} + \mathcal{O}(h^3),\\ &= -h^2(e \otimes f_z) + h^2\,(A \otimes f_z)(A^{-1}e \otimes I_N) + \mathcal{O}(h^3),\\ &= \mathcal{O}(h^3). \qquad (21)\end{aligned}$$

[5] See for instance Theorem 4.1 of [HW96]. The interested reader can find a complete proof in [CCM02a]

Now, we have that $y^* = y + \mathcal{O}(h)$, and taking (20), (21) into account, that $\partial y^*/\partial\lambda = \mathcal{O}(h^2)$. Hence, we get from $g(y^*) = 0$ that

$$
\begin{aligned}
0 = \frac{\partial}{\partial\lambda} g(y^*) \;&=\; g_y \frac{\partial y^*}{\partial\lambda} + \mathcal{O}(h^3),\\
\frac{\partial y^*}{\partial\lambda} &= h^2 f_z \left(-I_N + \frac{\partial\lambda^*}{\partial\lambda}\right) + h(b^T \otimes f_z)\frac{\partial\widehat{Z}}{\partial\lambda} + \mathcal{O}(h^3),\\
&= h^2 f_z \left(-R(\infty) I_N + \frac{\partial\lambda^*}{\partial\lambda}\right) + \mathcal{O}(h^3),
\end{aligned}
$$

which, taking the invertibility of $g_y f_z$ into account, leads to the estimates (13). □

Remark 1 The estimates (13) can be rewritten as

$$y^* = \Psi_h^{[1]}(y,\lambda) = \Psi_h^{[1]}(y,0) + \mathcal{O}(h^3||\lambda||), \tag{22}$$

$$\lambda^* = \Psi_h^{[2]}(y,\lambda) = (R(\infty) + \mathcal{O}(h))\lambda + \Psi_h^{[2]}(y,0), \tag{23}$$

where $\Psi_h(y,0)$ can now be interpreted as the composition of one step of the Runge-Kutta method followed by a projection step.

4 Convergence

Hereafter we will denote by φ_t the t-flow of the system (3). By definition of $\theta(y)$ and the flow φ_t of (3) we have that $\frac{d}{dt} g(\varphi_t(y)) \equiv 0$, and therefore $g(\varphi_t(y)) = g(y)$.

Lemma 1 *Let us consider the map $(y^*, \lambda^*) = \Psi_h(y,\lambda)$ corresponding to a given Runge-Kutta method $\mathcal{R} = (A, b)$.*

1. If $\mathcal{R}$ satisfies the conditions

$$C(q) \;:\; \sum_{j=1}^{s} a_{i,j} c_j^{k-1} = \frac{c_i^k}{k} \text{ for } k = 1,\dots,q \text{ and } i = 1,\dots,s,$$

then

$$
\begin{aligned}
\widehat{Y}_i - \varphi_{c_i h}(y) &= \mathcal{O}(h^{q+1} + h^3||\lambda||),\\
\widehat{Z}_i - \theta(\varphi_{c_i h}(y)) &= \mathcal{O}(h^q + h||\lambda||).
\end{aligned} \tag{24}
$$

2. If in addition $\mathcal{R}$ satisfies the conditions

$$B(q+1) \;:\; \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k} \text{ for } k = 1,\dots,q+1,$$

then

$$\Psi_h^{[1]}(y,\lambda) = \varphi_h(y) + \mathcal{O}(h^{q+2} + h^3||\lambda||),$$
$$\Psi_h^{[2]}(y,\lambda) = (R(\infty) + \mathcal{O}(h))\lambda + \mathcal{O}(h^{q-1}). \tag{25}$$

Proof According to Remark 1, it is sufficient to get estimates for $\Psi_h(y,0)$, i.e. for the Runge-Kutta method followed by a projection step. These estimates can be found for instance in Lemma 4.3. of [HLR89]. □

Theorem 2 *Let us assume that the Runge-Kutta scheme $\mathcal{R} = (A,b)$ satisfies $C(s)$ and $B(p)$ with $s \le p \le 2s$, and has non-zero abscissas. Given $(y,\lambda) \in \mathcal{M} \times B_N(0,\alpha)$ such that $\lambda = \mathcal{O}(h^{s-2})$, it holds for $|h| < \tilde{h}$ that*

$$\Psi_h^{[1]}(y,\lambda) = \varphi_h(y) + \mathcal{O}(h^{p+1} + h^3\|\lambda\|^2 + h^{s+2}\|\lambda\|).$$

Furthermore if f is linear in z, then

$$\Psi_h^{[1]}(y,\lambda) = \varphi_h(y) + \mathcal{O}(h^{p+1}).$$

Proof Let us consider the s-degree interpolating polynomial $\hat{u}(t)$ satisfying $u'(c_ih) = f(\widehat{Y}_i, \widehat{Z}_i)$, $1 \le i \le s$, and $\hat{u}(0) = \widehat{y}$. Since $C(s)$ and $B(p)$ are fulfilled with $p \ge s$, the Runge-Kutta method is a collocation polynomial (see for instance the proof of Theorem II.7.8 in [HNW93]), and therefore $\hat{u}(c_ih) = \widehat{Y}_i$, $1 \le i \le s$, $\hat{u}(h) = \widehat{y}^*$. The hypothesis $\lambda = \mathcal{O}(h^{s-2})$ and (24) (with $q = s$) imply that $\widehat{Y}_i = y(c_ih) + \mathcal{O}(h^{s+1})$. From that, it follows (see for instance the proof of Theorem II.7.10 in [HNW93]) that all the derivatives of $\hat{u}(t)$ are uniformly bounded on $[0,h]$ as $h \to 0$. We now consider $u(t) = \mathcal{P}_{\mathcal{M}}(\hat{u}(t))$. Clearly, $u(c_ih) = \hat{u}(c_ih) = \widehat{Y}_i$, $u(0) = y$, $u(h) = y^*$, and $u(t)$ is the solution of the initial value problem

$$u' = f(u, \theta(u)) + \delta(t), \quad u(0) = y,$$

where $\delta(t) = u'(t) - f(u(t), \theta(u(t)))$. Then, we can estimate $y^* - \varphi_h(y) = u(h) - \varphi_h(y)$ by means of the nonlinear variation-of-constants formula (see Theorem I.14.5 of [HNW93]),

$$u(h) - \varphi_h(y) = \int_0^h R(t)\delta(t)dt \text{ where } R(t) = \frac{\partial}{\partial y}\varphi_{h-t}(u(t)). \tag{26}$$

Since the assumption $B(p)$ means that the quadrature formula with c_i as nodes and b_i as weights is of order p, we arrive at

$$y^* - \varphi_h(y) = h\sum_{i=1}^{s} b_i R(c_ih)\delta(c_ih) + \mathcal{O}(h^{p+1}).$$

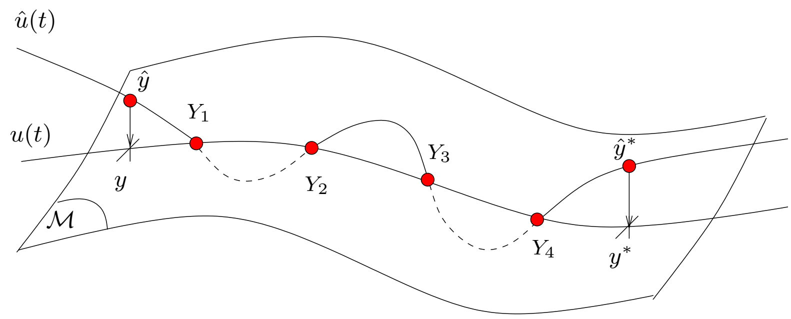

Fig. 1. Polynomial $\hat{u}(t)$ and its projection $u(t)$.

Note that, for this to be correct, we need that the derivatives of $\delta(t)$ are uniformly bounded on $[0,h]$ as $h \to 0$. This is a consequence of the uniform bounds of the derivatives of $\hat{u}(t)$ and the smoothness of the functions f, g, and $\mathcal{P}_\mathcal{M}$.

It follows from the definition of the projection $\mathcal{P}_\mathcal{M}(y)$ that, for $y \in \mathcal{M}$, we have

$$\frac{\partial \mathcal{P}_\mathcal{M}}{\partial y}(y) = P(y), \text{where} \quad P(y) \;=\; I - f_z(g_y f_z)^{-1} g_y(y, \theta(y)).$$

Differentiating the equality $u(t) = \mathcal{P}_\mathcal{M}(\hat{u}(t))$ with respect to t and evaluating at $t = c_i h$, we obtain

$$u'(c_i h) = \frac{\partial \mathcal{P}_\mathcal{M}}{\partial y}(\hat{u}(c_i h))\hat{u}'(c_i h) = P(\widehat{Y}_i) f(\widehat{Y}_i, \widehat{Z}_i), \quad 1 \le i \le s. \tag{27}$$

It is straightforward to check the identities $P(y)f(y,\theta(y)) \equiv f(y,\theta(y))$ and $P(y)f_z(y,\theta(y)) \equiv 0$, which together with (27) imply that

$$\delta(c_i h) = P(\widehat{Y}_i)(f(\widehat{Y}_i, \widehat{Z}_i) - f(\widehat{Y}_i, \theta(\widehat{Y}_i)) - f_z(\widehat{Y}_i, \theta(\widehat{Y}_i))(\widehat{Z}_i - \theta(\widehat{Y}_i))).$$

This identically vanishes if f is linear in z, and is $\mathcal{O}(||\widehat{Z}_i - \theta(\widehat{Y}_i)||^2)$ otherwise. The required result for general f follows from application of Lemma 1 with $q = s$. $\square$

From these local error estimates one can now derive global convergence results in a straigthforward way : it must be indeed emphasized once again that φ_t is defined as the flow of an ODE so that familiar techniques for studying ODEs apply.

Remark 2 Note that the order of Ψ_h can also be interpreted as the *effective order* of Φ_h, as introduced in [But69], where the role of the finishing procedure is played by the projection step.

5 Backward error analysis for symmetric methods

The reversibility of the numerical method is a key ingredient to backward analysis, so that we now investigate this question for Ψ_h. To show that Ψ_h is Θ-reversible with

$$\Theta = \begin{bmatrix} \rho & 0 \\ 0 & -I_M \end{bmatrix},$$

we prove, following [HLW02], that, on one hand, $\Psi_h \circ \Theta = \Theta \circ \Psi_{-h}$ and on the other hand, Ψ_h is time-reversible or symmetric, i.e. satisfies $\Psi_{-h} = \Psi_h^{-1}$.

We first state without proof the following lemma and then proceed with establishing the two equalities for Ψ_h :

Lemma 2 *Suppose that the differential-algebraic system (1) satisfies the reversibility conditions (2). Then the following equalities hold :*

$$\begin{aligned} \forall\,(y,z) \in \mathbb{R}^M \times \mathbb{R}^N,\ & f_z(\rho y, z) = -\rho f_z(y,z), \\ & g_y(\rho y)\,\rho = \tilde{\rho}\, g_y(y), \\ \forall\, y \in \mathcal{V} \text{ such that } \rho y \in \mathcal{V},\ & \theta(\rho y) = \theta(y). \end{aligned}$$

Theorem 3 *Suppose that the differential-algebraic system (1) satisfies the reversibility conditions (2). For any $|h| \le \tilde{h}$, we have*

$$\begin{aligned} \forall\,(y,\lambda) \in \mathcal{M} \times B_N(0,\alpha),\ & \Psi_h^{[1]}(\rho y, -\lambda) = \rho(\Psi_{-h}^{[1]}(y,\lambda)), \\ & \Psi_h^{[2]}(\rho y, -\lambda) = -(\Psi_{-h}^{[2]}(y,\lambda)). \end{aligned}$$

Proof Going back to the definition of Ψ_h, and replacing y, λ and h respectively by ρy, $-\lambda$ and $-h$ in equation (9), we get

$$\begin{aligned} \rho y - (-h)^2 f_z(\rho y, \theta(\rho y))(-\lambda) &= \rho y + h^2(\rho\, f_z(y,\theta(y)))(-\lambda), \\ &= \rho\widehat{y}, \end{aligned}$$

where $\widehat{y}$ is given by (9). Now, since $\Phi_{-h}(\rho\widehat{y}) = \rho\Phi_h(\widehat{y})$, the values $y^{**} = \Psi_{-h}^{[1]}(\rho y, -\lambda)$ and $\lambda^{**} = \Psi_{-h}^{[2]}(\rho y, -\lambda)$ are solution of the system

$$\begin{aligned} y^{**} &= \rho\widehat{y}^{*} + (-h)^2 f_z(y^{**}, \theta(y^{**}))\lambda^{**}, \\ 0 &= g(y^{**}) - g(\rho y), \end{aligned}$$

and must, by uniqueness, coincide with ρy^* and $-\lambda^*$. $\square$

Theorem 4 *Suppose that the Runge-Kutta method $\mathcal{R} = (A,b)$ is* symmetric. *For any $|h| < \tilde{h}$ and for all $y \in \mathcal{M}$ and all $\lambda \in B_N(0,\alpha)$ such that $\Psi_h^{[2]}(y,\lambda) \in B_N(0,\alpha)$, the following equality holds :*

$$\Psi_{-h}(\Psi_h(y,\lambda)) = (y,\lambda).$$

Proof $\mathcal{R}$ being symmetric, the result follows immediately from the fact that the mapping $\mathcal{P}$ (see Diagram (12)) remains the same for h and $-h$. □

6 Symmetrically projected methods

In this section we consider the adaptation of *symmetrically projected methods*, as proposed in [Hai00, Hai01] for ODEs, to index-2 DAEs. Given a p^{th}-order *symmetric* Runge-Kutta method $\mathcal{R} = (A, b)$, the symmetric projection is a procedure aimed at preserving *reversibility* while assuring at the same time that the numerical approximation lies on the manifold $\mathcal{M}$. It proceeds in three steps (see Figure 2) and it is a map S_h from $\mathcal{M}$ to itself :

$$\begin{cases} \widehat{y} \;= y - h^2 f_z(y, \theta(y))\lambda, \\ \widehat{y}^* = \Phi_h(\widehat{y}), \\ y^* = \widehat{y}^* - h^2 R(\infty) f_z(y^*, \theta(y^*))\lambda, \; g(y^*) = 0, \end{cases} \tag{28}$$

where $R(\infty) = 1 - b^T A^{-1} e$ is the value of the stability function of $\mathcal{R}$ at infinity. The reversibility of the procedure may be obtained as

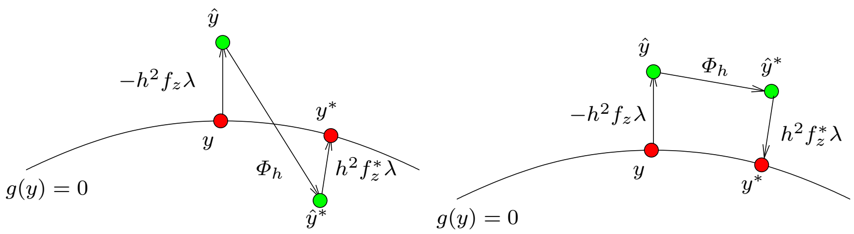

Fig. 2. Symmetric projection with $R(\infty) = 1$ (left) and $R(\infty) = -1$ (right)

a by-product of the reversibility of Ψ_h. Equations (28) can indeed be rewritten as

$$(y^*, \lambda^*) = \Psi_h(y, \lambda), \tag{29}$$
$$\lambda^* = -R(\infty)\lambda. \tag{30}$$

The existence of (y^*, λ) satisfying equations (28) or equivalently (29) is a simple consequence of the smoothness of Ψ_h. Actually, equations (28) define for sufficiently small $|h| \leq \bar{h}$ a one-step method S_h :

Definition 2 *We define the map S_h parametrised by h*

$$\begin{aligned} S_h \;&: \mathcal{M} \to \mathcal{M} \\ y \;&\mapsto y^* = S_h(y) \end{aligned}$$

by the equations

$$\begin{aligned} 0 &= R(\infty)\lambda + \Psi_h^{[2]}(y,\lambda), \\ y^* &= \Psi_h^{[1]}(y,\lambda). \end{aligned}$$

Theorem 5 *For sufficiently small $|h| \le \bar{h}$, the map S_h is* smooth, *also with respect to h. Moreover, it is reversible and of order p if $C(s)$ and $B(p)$ are satisfied with $s \le p \le 2s$.*

Proof We look for a zero of the function

$$K(\lambda) = R(\infty)\lambda + \Psi_h^{[2]}(y,\lambda).$$

Since $|R(\infty)| = 1$, it can be checked that

$$\begin{aligned} \frac{\partial K}{\partial \lambda} &= R(\infty) I_N + \frac{\partial \Psi_h^{[2]}}{\partial \lambda}, \\ &= 2R(\infty) I_N + \mathcal{O}(h), \end{aligned}$$

is invertible for sufficiently small $|h|$. Now, for $h = 0$, $K(0) = 0$, so that the *Implicit Function Theorem* locally defines the solution λ of $K(\lambda) = 0$ as a smooth function of y and h. Provided $0 < \bar{h} < \tilde{h}$ is small enough, S_h can thus be smoothly extended over $\mathcal{M}$ for all $|h| \le \bar{h}$.

In order to prove that $y^* = S_h(y)$ is an approximation of order p to the solution, it suffices to notice that $\lambda = \mathcal{O}(h)$. Inserted in the estimate (25), this gives $\lambda = \mathcal{O}(h^{s-1})$ and thus

$$y^* = \varphi_h(y) + \mathcal{O}(h^{\min(p+1, 3+2(s-1))})$$

due to Theorem 2. □

Remark 3 The one-step method S_h being reversible and smoothly defined on $\mathcal{M}$, a backward error analysis can be carried out just as in previous section. The main difference is now that the corresponding *modified* equation is an equation for y only, with first term $f(y, \theta(y))$.

7 Numerical experiments

In this section, we compare the long-term behavior of post-projection and symmetric projection for the Gauss methods. Since the two techniques give similar results for the pendulum problem, we present numerical experiments for the "double pendulum" whose equations of motion may be obtained from the Hamiltonian function

$$H(q,p) = \frac{1}{2}(p_{x_1}^2 + p_{z_1}^2) + \frac{1}{2}(p_{x_2}^2 + p_{z_2}^2) + z_1 + z_2$$

and the two constraints

$$0 = x_1^2 + z_1^2 - 1, \tag{31}$$
$$0 = (x_2 - x_1)^2 + (z_2 - z_1)^2 - 1, \tag{32}$$

where we have assumed that the two bodies have unit mass, that the gravitational constant is 1 and that the two rods are massless and of length 1. In this form, the problem is of index 3 and the constraints (31) and (32) have to be replaced by their first derivatives with respect to time to get the index-2 formulation that we use for our numerical tests. In our tests, we have used the following initial values

$$(x_1, z_1, x_2, z_2, p_{x_1}, p_{z_1}, p_{x_2}, p_{z_2}) = (0.5, -\sqrt{0.75}, 0, -2\sqrt{0.75}, 0, 0, 0, 0)$$

corresponding to $H = -3\sqrt{0.75}$, and a stepsize $h = 0.1$.

Comparing the results of the two figures, we can make two main observations :

- the Hamiltonian is much better preserved by symmetrically projected methods than by post-projected methods. Symmetrically projected methods leaves the numerical solution on the constraint manifold, while post-projected methods embeds a drift which becomes dominant in the error.
- the three-stage post-projected method behaves much better than the two-stage method, and more generally, as additional experiments confirm it, methods with an odd number of stages behave better than methods with an even number of stages. This is of course related to the value of the stability function at infinity. Although we do not have a definite explanation of this phenomenon at this stage, we suspect that the modified equation governing the evolution of λ is stable only for $R(\infty) = 1$.

Acknowledgments

The authors are very grateful to one of the referees for pointing out a flaw in the numerical experiments and for his numerous suggestions.

References

[But69] J.C. Butcher. The effective order of Runge-Kutta methods. *Lecture Notes in Mathematics*, vol. 109, pages 133–139, 1969. Conference on the numerical solution of differential equations.

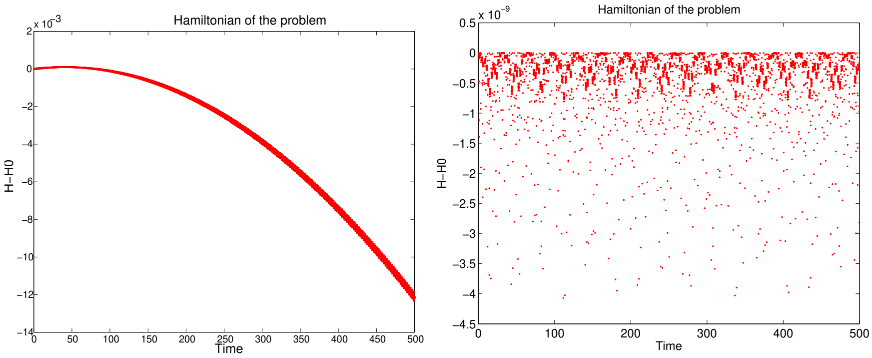

Fig. 3. Evolution of the Hamiltonian for the 2-stage (left) and the 3-stage (right) Gauss methods with post-projection

[CCM02a] R.P.K. Chan, P. Chartier, and A. Murua. A new convergence analysis of Runge-Kutta methods for index-2 differential-algebraic equations. Technical report, INRIA, 2002.

[CCM02b] R.P.K. Chan, P. Chartier, and A. Murua. Post-projected Runge-Kutta methods for index-2 differential-algebraic equations. *Journal of Applied Numerical Mathematics*, 42:77–94, 2002.

[GSH99] O. Gonzalez, A.M. Stuart, and D.J. Higham. Qualitative properties of modified equations. *IMA Journal of Numerical Analysis*, 19:169–190, 1999.

[Hai00] E. Hairer. Symmetric projection methods for differential equations on manifold. *BIT*, 40(4):726–734, 2000.

[Hai01] E. Hairer. Geometric integration of ordinary differential equations on manifold. *BIT*, 41:996–1007,2001.

[HLR89] E. Hairer, Ch. Lubich, and M. Roche. *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*. Springer-Verlag, 1989. Lecture Notes in Mathematics 1409.

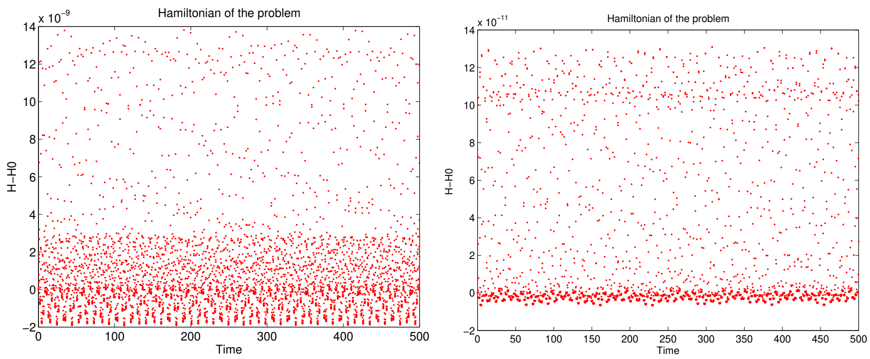

Fig. 4. Evolution of the Hamiltonian for the 2-stage (left) and 3-stage (right) symmetrically projected Gauss methods

[HLW02] E. Hairer, Chr. Lubich, and G. Wanner. *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, 2002.

[HNW93] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations, Nonstiff Problems*. Springer-Verlag, second revised edition, 1993. Volume 1.

[HW96] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II, Stiff Problems and Differential-Algebraic Problems*. Springer-Verlag, second revised edition, 1996. Volume 2.

[Rei97] S. Reich. On higher-order semi-explicit symplectic partitioned Runge-Kutta methods for constrained Hamiltonian systems. *Numerische Mathematik*, 76:231–247, 1997.

[Sto88] D. Stoffer. On reversible and canonical integration methods. Technical Report 88-05, ETH-Zürich, 1988.