BIT
2001, Vol. 41, No. 3, pp. 599–628
0006-3835/01/4103-0599 $16.00

ON THE NUMERICAL SOLUTION OF INVOLUTIVE ORDINARY DIFFERENTIAL SYSTEMS: HIGHER ORDER METHODS *

J. TUOMELA[1] and T. ARPONEN[2]

[1] *Department of Mathematics, University of Joensuu, P.O. Box 111 FI-80101 Joensuu, Finland. email: jukka.tuomela@joensuu.fi*

[2] *Institute of Mathematics, Helsinki University of Technology, P.O. Box 1100 FI-02015 HUT, Finland. email: teijo.arponen@hut.fi*

Abstract.

We analyse some Taylor and Runge–Kutta type methods for computing one-dimensional integral manifolds, i.e. solutions to ODEs and DAEs. The distribution defining the solutions is taken to be defined only on the relevant manifold and hence all the intermediate points occuring in the computations are projected orthogonally to the manifold. We analyse the order of such methods, and somewhat surprisingly there does not appear any new order conditions for the Runge–Kutta methods in our context, at least up to order 4. The analysis shows that some terms appearing in the error expansions can be quite naturally expressed in terms of standard notions of Riemannian geometry. The numerical examples show that the methods work reliably and moreover produce qualitatively correct results for Hamiltonian systems although the methods are not symplectic.

AMS subject classification: primary 34A26, 65L05, secondary 35N10, 58F40.

Key words: Differential algebraic equations, integral manifolds, Runge–Kutta methods.

1 Introduction.

We continue the work started in [13] where we formulated differential systems (ODEs, DAEs, etc) using jet bundles and showed that this approach was useful and interesting also from the numerical point of view. Here we shall pursue this topic further by studying Runge–Kutta and Taylor type methods which are relevant in this context. Formulating the problem in jet bundles leads to computing one-dimensional integral manifolds (or curves) of a distribution (or a vector field) on a manifold which is embedded in some Euclidean space. Note that one may quite naturally arrive at this formulation also without using jets: for instance in Hamiltonian problems the relevant vector field restricts to the manifold where the Hamiltonian is constant. Hence the results obtained below can be interesting also in some numerical problems where the jets are not explicitly used.

We shall start by analysing some Taylor type methods and compute explicitly the terms needed to get methods of orders up to three. It is seen that the

*Received June 1999. Revised January 2001. Communicated by Gustaf Söderlind.

information required can quite naturally be expressed in terms of standard concepts of Riemannian geometry. Since we expect that Runge–Kutta methods are more useful in practice we devote more time to them. Now there are many ways one could try to formulate Runge–Kutta methods in the present context. We take the point of view that the vector field or distribution is given only on the manifold, and hence all the intermediate results have to be projected back to the manifold. Moreover the projection is required to be orthogonal. Then the vectors are combined in ordinary fashion, interpreting them to be vectors in the ambient space. Proceeding in this fashion in the construction of the methods, we find somewhat surprisingly that there are *no* new order conditions, at least for methods of order up to four. The question naturally arises if this property holds in general. Unfortunately the proofs do not admit an immediate generalisation which is required in this general question.

Another way to implement Runge–Kutta methods would be to use parallel translation in the addition of vectors. This would be quite difficult from the practical point of view, because in that case one would have to construct explicitly coordinates for the manifold and then solve (numerically) the differential equations which give the parallel translation. Although this appears numerically unattractive, it might still be of theoretical interest to analyse this situation more carefully.

One could also try to analyse multistep methods in a similar fashion, and the computation of the order conditions would proceed much in the same way as for Runge–Kutta methods. However, the stability analysis would be quite difficult and would require considerations that are beyond the scope of the present paper. Perhaps more importantly, *linear* multistep methods may appear a bit dubious at the outset: when the computations are done on manifolds, one may wonder if the past information is as useful as it is in the linear spaces.

In [13] the background and motivations for using jets was explained in great detail with extensive references to relevant literature. Hence in the present article we simply recall the notations and refer to [13] for more details in order to minimize repetitions.

Since there are no new conditions for Runge–Kutta methods (at least for methods of order up to four), and the ambient space is used in the combination of vectors, the implementation of these methods is quite straightforward, at least in principle. However, trying to produce an efficient code is a non-trivial task; this, together with comparisons to other methods dealing with DAEs, will be considered elsewhere. The purpose of the present article is to demonstrate that it is possible to have high order methods in our context and that such methods work reliably. In particular it is seen that overdetermined systems are not intrinsically harder than determined ones. Note also that the methods used in the computations are explicit, contrary to other existing methods for these kind of systems.

In the final section we then present some numerical results obtained. The computations lend some support to the conjecture that there would not be any new order conditions even for methods of order greater than four. In the case of Hamiltonian systems it is seen that the results are qualitatively correct over quite

long time intervals although our method is not symplectic. Moreover, symplectic methods usually perform better when a fixed step-size is used. In our method changing the step-size has no such effect. In the final example we also show how it is possible to use ideal theory to remove constraint singularities from the system.

This article is based on reports [12] and [14] where some tedious computations can be found which are suppressed here.

2 Basic tools.

We recall briefly the main notions that are needed. For more details on standard differential geometry we refer to [11] and on jets to [9]. All maps and manifolds are assumed to be smooth, i.e. infinitely differentiable. All analysis is local, hence various maps and manifolds need to be smooth or defined only in some appropriate subsets. To simplify the notation these subsets are not indicated. Moreover, if M is a submanifold of $\tilde{M}$, then objects defined on M can be taken to be defined on $\tilde{M}$ without writing explicitly the inclusion map.

2.1 Riemannian geometry.

The jth differential of a map $f : \mathbb{R}^m \mapsto \mathbb{R}^k$ is denoted by $d^j f$ and its value at p by $d^j f_p$. Let M be a manifold. The set of maps $M \to \mathbb{R}$ is denoted by $C^\infty(M)$, the tangent bundle of M by TM, and the tangent space at $p \in M$ by TM_p. A distribution $\mathcal{D}$ is a map that associates to each point $p \in M$ a subspace $\mathcal{D}_p$ of TM_p. If D is one-dimensional in some open set Ω and $p \in \Omega$, then in a sufficiently small neighborhood of p there exists a vector field (actually two of them) V such that $|V_p| = 1$ and $V_p \in \mathcal{D}_p$. Such V is said to be associated to D. If $\mathcal{E}$ is a bundle, then the set of its (local) sections is denoted by $\Gamma(\mathcal{E})$. Tangent vectors can be identified with differential operators, thus if $f \in C^\infty(M)$ and $X \in \Gamma(TM)$, then $X(f) \in C^\infty(M)$.

Let M be a submanifold of $\mathbb{R}^m$ for some m with standard Riemannian metric. We give M the Riemannian metric induced by this embedding. Recall that Riemannian metric is a positive definite bilinear map on the tangent space TM_p which varies differentiably with p. It is denoted by $\langle \cdot, \cdot \rangle$ and the same notation is used for the standard inner product in $\mathbb{R}^m$. The normal bundle of M with respect to $\mathbb{R}^m$ is denoted by NM and the normal space at p by NM_p. Since $(T\mathbb{R}^m)_p = TM_p \oplus NM_p$ we have the orthogonal projections $\pi_t : (T\mathbb{R}^m)_p \mapsto TM_p$ and $\pi_n : (T\mathbb{R}^m)_p \mapsto NM_p$. Recall that sections of TM and NM can locally be extended to sections of $T\mathbb{R}^m$. These extensions will be denoted by the same symbol as the original sections.

The unique symmetric connection on M compatible with metric is denoted by ∇. There are many equivalent definitions of a connection; for our purposes it is convenient to regard ∇ as a map $\Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$, denoted by $(X, Y) \mapsto \nabla_X Y$, which satisfies the following conditions:

- $\nabla_X Y$ is $\mathbb{R}$-linear in Y,
- $\nabla_X Y$ is $C^\infty(M)$-linear in X,

- $\nabla_X(fY) = f\,\nabla_X Y + X(f)\,Y$ (Leibniz rule).

Fixing $X \in \Gamma(TM)$ we may define a map $\nabla_X : \Gamma(TM) \to \Gamma(TM)$ by $Y \mapsto \nabla_X Y$; this is called the covariant derivative of Y with respect to X. In standard Euclidean space it is just the directional derivative. Let us further recall that the bracket is a map $[\cdot,\cdot] : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$, defined by $[X,Y](f) = X(Y(f)) - Y(X(f))$, where $f \in C^\infty(M)$. The same notation is used when the bracket is interpreted in $\mathbb{R}^m$. In terms of the bracket, the symmetry of the connection means that

$$\nabla_X Y - \nabla_Y X = [X,Y]. \tag{2.1}$$

The compatibility of the connection with the metric is equivalent to

$$X\langle Y,Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle. \tag{2.2}$$

Let us finally recall the following basic property. Let $p \in M$ and $X\,, Y \in \Gamma(TM)$. Then

$$dY(X_p) = \pi_t\big(dY(X_p)\big) + \pi_n\big(dY(X_p)\big) = \nabla_{X_p} Y + S(X_p, Y_p), \tag{2.3}$$

where S is a symmetric tensor, called the second fundamental tensor.

2.2 Differential systems in jet spaces.

Here we simply give the basic definitions and refer to [13] for a discussion and motivation of these concepts. Let $\pi : \mathcal{E} \to \mathcal{B}$ be a bundle and let $J_q(\mathcal{E})$ be the bundle of q-jets of $\mathcal{E}$.

Definition 2.1. *A (partial) differential system (or equation) of order q on $\mathcal{E}$ is a submanifold $\mathcal{R}_q$ of $J_q(\mathcal{E})$.*

Let $\mathcal{E} = \mathbb{R} \times \mathbb{R}^n$ and let us denote the coordinates of $J_q(\mathcal{E})$ by $(x, y^1, \dots, y^n, y^1_1, \dots, y^n_q)$. Let us define the one forms

$$\alpha^i_j = dy^i_{j-1} - y^i_j dx, \quad i = 1, \dots, n, \quad j = 1, \dots, q. \tag{2.4}$$

Let $p \in J_q(\mathcal{E})$ and $v_p \in (TJ_q(\mathcal{E}))_p$ and let us further set

$$\begin{aligned} C_p &= \big\{ v_p \in (TJ_q(\mathcal{E}))_p \,\big|\, \alpha^i_j(v_p) = 0 \big\}, \\ \mathcal{D}_p &= (T\mathcal{R}_q)_p \cap C_p. \end{aligned} \tag{2.5}$$

Now we can define the solutions as follows.

Definition 2.2. *Let $\mathcal{R}_q \subset J_q(\mathcal{E})$ be involutive and suppose that the distribution $\mathcal{D}$ defined in* (2.5) *is one-dimensional. A solution of $\mathcal{R}_q$ is an integral manifold of $\mathcal{D}$.*

We must refer to [13] for an extensive discussion and further references on involutivity. Intuitively a system is involutive if we cannot get new equations of order q or less by differentiating the equations and eliminating the higher derivatives.

Suppose that we are given a system of k ordinary differential equations of qth order

$$f(x, y, y_1, \dots, y_q) = 0. \tag{2.6}$$

In terms of coordinates f can be taken to be a map $\mathbb{R}^{(n+1)q+1} \to \mathbb{R}^k$ and so the zero set of the above equation defines a certain submanifold of $J_q(\mathcal{E})$ which we denote by $\mathcal{R}_q$. Let us assume that $\mathcal{R}_q$ is involutive and the corresponding distribution $\mathcal{D}$ one-dimensional. Since one-dimensional distributions always have integral manifolds, there always exists a solution to our problem in these circumstances. In the following sections we discuss how to compute these solutions.

All the methods discussed are one-step methods. Let us recall the following basic property of these methods [3, 7]: if the local error is $O(h^{r+1})$ then the global error is $O(h^r)$ for sufficiently small h.

3 Taylor type methods.

Let $k < m$ and $f : \mathbb{R}^m \mapsto \mathbb{R}^k$ and $M := f^{-1}(0) \subset \mathbb{R}^m$. Let $\mathcal{D}$ be a smooth one-dimensional distribution on M and V a smooth vector field associated to it, at a given point $p \in M$. We would like to compute the integral manifold that passes through p.

Let $p \in M$ be the current point and $c : \mathbb{R} \mapsto M$ be the integral curve of the associated vector field with $c(0) = p$ and $c'(0) = V_p$. Note that c is parametrized by arclength. Let $w \in (T\mathbb{R}^m)_p$; then an approximation to $c(h)$ is obtained by computing the solution q of the following system:

$$\begin{cases} q + (df_q)^t \mu = p + hw, \\ f(q) = 0, \end{cases} \tag{3.1}$$

where $\mu \in \mathbb{R}^k$. Note that the solution exists for h sufficiently small. It will be convenient to define a curve $\tilde{c}(s) =$ solution q of (3.1) with $h = s$. Evidently $c(0) = \tilde{c}(0) = p$, and we would like to measure the difference $c(h) - \tilde{c}(h)$ for small h and to choose w in such a way that this difference is as small as possible. Let us start with simple

LEMMA 3.1. *Let $\tilde{c}$ and c be as above; if $\tilde{c}^{(k)}(0) = c^{(k)}(0)$ for $1 \le k < n$, then $\tilde{c}^{(n)}(0) - c^{(n)}(0) \in TM_p$.*

PROOF. The curve c satisfies identically $f \circ c = 0$. Hence

$$\frac{d^n}{ds^n}\big(f \circ c\big) = df\, c^{(n)} + \text{terms with lower order derivatives of } c = 0.$$

The same holds for $\tilde{c}$, so the hypothesis implies that $df_p\big(\tilde{c}^{(n)}(0) - c^{(n)}(0)\big) = 0$. □

We then use expansions to determine 'good' directions. Let us consider the system

$$\begin{cases} q + (df_q)^t \mu = p + \sum_{k=1}^n w^k h^k, \\ f(q) = 0. \end{cases} \tag{3.2}$$

We would like to find $w^k \in (T\mathbb{R}^m)_p$ in such a way that $c(h) - q = O(h^{n+1})$. The next result shows that good directions of arbitrary order exist, even if w^k are restricted to TM_p.

THEOREM 3.2. *For any n there are vectors $w^k \in TM_p$, $1 \le k \le n$, such that if q is a solution of (3.2), then $c(h) - q = O(h^{n+1})$.*

PROOF. The case $n = 1$ follows by choosing $w^1 = V_p$. Let us suppose there are $w^k \in TM_p$, $1 \le k < n$, such that $c(h) - p^* = O(h^n)$, where

$$\begin{cases} p^* + (df_{p^*})^t \mu = p + \sum_{k=1}^{n-1} w^k h^k, \\ f(p^*) = 0. \end{cases}$$

Let $p^* = p + p^1 h + p^2 h^2 + \cdots$; then by Lemma 3.1 $\frac{1}{n!} c^{(n)}(0) - p^n \in TM_p$. Put $w^n = \frac{1}{n!} c^{(n)}(0) - p^n$ and let q be a solution of (3.2) with this w^n. Let $q = p + q^1 h + q^2 h^2 + \cdots$; evidently $p^k = q^k$ for $1 \le k < n$, so we have to show that $q^n = \frac{1}{n!} c^{(n)}(0)$.

Now expanding the second equation in (3.2) we get $df_p(q^n - p^n) = 0$. Hence $\pi_n(q^n) = \pi_n(p^n) = \pi_n(\frac{1}{n!} c^{(n)}(0))$. Then expanding the first equation we get

$$q^n + (df_p)^t \mu^n + b^n = w^n,$$

where b^n contains the terms with μ^k, $k < n$. Note that b^n does not contain q^n (since $\mu = O(h^2)$), and consequently b^n is the same for p^* and q. Now

$$\pi_t(q^n) = \pi_t(w^n) - \pi_t(b^n) = \pi_t\big(\tfrac{1}{n!} c^{(n)}(0) - p^n\big) + \pi_t(p^n) = \pi_t\big(\tfrac{1}{n!} c^{(n)}(0)\big).$$

□

Let us compute the second order terms in (3.2) with $w^1 = V_p$ and $w^2 = 0$. We get the system

$$\begin{cases} q^2 + (df)^t \mu^2 = 0, \\ df\, q^2 + \frac{1}{2}\, d^2 f(V_p, V_p) = 0. \end{cases}$$

To proceed in the computation let us introduce some convenient notations. Let $B = df(df)^t$ and let S_V (resp. S_∇) denote the section of the normal bundle defined by $S(V,V)$ (resp. $S(V, \nabla_V V)$). From now on we shall also drop the subscript p in formulas like $\nabla_{V_p} V$ when the meaning is clear from the context. We shall need the following lemma:

LEMMA 3.3. *Let V, $Y \in \Gamma(TM)$; then*

$$S(V,Y) = -(df)^t B^{-1} d^2 f(V,Y).$$

PROOF. Let $c : \mathbb{R} \mapsto M$ with $c(0) = p$ and $c'(0) = V_p$. Then $S(V_p, V_p)$ is given by the above formula, since $S(V_p, V_p) = \pi_n(c''(0))$. The general result follows by bilinearity and symmetry of S. □

Hence using Lemma 3.3 we compute that $q^2 = \frac{1}{2}\, S(V,V)$ which combined with $c''(0) = dV\, V$ and (2.3) implies that

$$\tfrac{1}{2}\, c''(0) - q^2 = \tfrac{1}{2}\, \nabla_V V.$$

Hence by the proof of the previous theorem we obtain at once

COROLLARY 3.4. *If q is the solution of* (3.2) *with $w^1 = V$ and $w^2 = \frac{1}{2}\nabla_V V$, then $c(h) - q = O(h^3)$.*

Note that the correction terms w^1 and w^2 are just what one would expect them to be by the classical theory. In the computations that follow the formula (2.3) and Lemma 3.3 are used very often. Note also that $\pi_n = (df)^t B^{-1} df$. Let us then compute the next term.

PROPOSITION 3.5. *If q is the solution of* (3.2) *with $w^1 = V$, $w^2 = \frac{1}{2}\nabla_V V$ and*

$$w^3 = \tfrac{1}{6}\nabla_V(\nabla_V V) + \tfrac{1}{3}\pi_t\big((dV)^t S_V\big),$$

then $c(h) - q = O(h^4)$.

PROOF. Expanding the system (3.2) we first compute that $q^2 = \frac{1}{2} dV\, V$ and $B\mu^2 = -\frac{1}{2} df\, S_V$. Then the third order terms are obtained from

$$\begin{cases} q^3 + (df)^t\mu^3 - \frac{1}{2} d^2 f(V,\,\cdot\,)B^{-1} df S_V = w^3, \\ df\, q^3 + \frac{1}{2} d^2 f(dV\, V, V) + \frac{1}{6} d^3 f(V,V,V) = 0. \end{cases} \tag{3.3}$$

By Theorem 3.2, it is sufficient to choose $w^3 \in TM_p$ in such a way that $\pi_t(q^3) = \pi_t(\frac{1}{6} c'''(0))$. Hence we need not compute μ^3 at all and can ignore the second equation in (3.3). Using Lemma 8.1 we get from the first equation

$$\pi_t(q^3) = w^3 - \tfrac{1}{2}\pi_t\big((dV)^t S_V\big).$$

Now combining Lemmas 8.2 and 8.3 leads to the result. □

There is now a non-classical correction term $X_p = \pi_t\big((dV)^t S_V\big)$. Let us then give a more geometric characterization of this term.

LEMMA 3.6. *Let $\{z^k\}$ be an orthonormal basis of TM_p. Then*

$$X_p = \sum_k \big\langle S_V, S(V, z^k)\big\rangle z^k. \tag{3.4}$$

PROOF. If $z^k \in TM_p$, then

$$\big\langle z^k, (dV)^t S_V\big\rangle = \big\langle dV\, z^k, S_V\big\rangle = \big\langle S(V, z^k), S_V\big\rangle.$$

□

Note that X_p depends only on V at p. Now Theorem 3.2 says that we need only tangential directions to get arbitrarily high order. However, it may still be useful to consider also normal directions. Let us start with the following simple observation:

PROPOSITION 3.7. *If q is the solution of* (3.2), *then*

$$\pi_t(w^1) = V_p \quad\Longleftrightarrow\quad c(h) - q = O(h^2).$$

Hence the normal component of w^1 has no effect for a first order method. However, it obviously affects higher order error terms. Could we choose $\pi_n(w^1)$ in such a way that $c(h) - q = O(h^3)$? Unfortunately we have the following:

LEMMA 3.8. *Let q be the solution of* (3.1) *with $w = V_p + Y_p$, where $Y_p \in NM_p$. If $\dim(TM_p) > \dim(NM_p)$, then in general it is impossible to choose Y_p such that $c(h) - q = O(h^3)$.*

PROOF. See [12]. □

Of course, nothing guarantees third order local error, even if $\dim(TM_p) \leq \dim(NM_p)$. In spite of the above result, the normal directions *are* useful.

PROPOSITION 3.9. *If q is the solution of* (3.2) *with $w^1 = V_p$, $w^2 = \frac{1}{2}(\nabla_{V_p} V)_p + \frac{1}{3} S_V$ and $w^3 = \frac{1}{6}\big(\nabla_{V_p}(\nabla_{V_p} V)\big)_p$, then $c(h) - q = O(h^4)$.*

PROOF. From

$$\begin{aligned} &q^2 + (df)^t \mu^2 = \tfrac{1}{2}(\nabla_{V_p} V)_p + \tfrac{1}{3} S_V, \\ &df\, q^2 + \tfrac{1}{2} d^2 f(V,V) = 0, \end{aligned}$$

we get $B\mu^2 = -\frac{1}{6} df S_V$ and $q^2 = \frac{1}{2} dVV$. Using this μ^2 in (3.3) gives

$$\pi_t(q^3) = \tfrac{1}{6} \nabla_V(\nabla_V V) - \tfrac{1}{6} \pi_t\big((dV)^t S_V\big),$$

which yields the result. □

Note that computing just S_V is much easier than computing X_p in (3.4). It is seen that the vectors w^i for $i = 1, 2, 3$ admit direct interpretations in terms of standard operations in Riemannian geometry. Characterizing higher order terms in this way seems to be more difficult. However, our main interest is in Runge–Kutta methods, since we expect that they are more useful in practice, and consequently we now turn our attention to them.

4 Runge–Kutta type methods.

4.1 Explicit methods.

We have already seen that one obtains a first order method by taking an Euler step along V_p, i.e. if q is a solution of

$$\begin{cases} q + (df_q)^t \mu = p + hV_p, \\ f(q) = 0, \end{cases}$$

then $c(h) - q = O(h^2)$. Now let us try to construct an explicit two stage Runge–Kutta scheme whose local error is $O(h^3)$. In our context this can be formulated in the following way:

$$\begin{cases} \tilde{q} + (df_{\tilde{q}})^t \tilde{\mu} = p + ha_{21} V_p, \\ f(\tilde{q}) = 0, \\ q + (df_q)^t \mu = p + h\big(b_1 V_p + b_2 V_{\tilde{q}}\big), \\ f(q) = 0. \end{cases} \tag{4.1}$$

Here we use trivial parallel translation in the ambient space to add vectors with different base points. Hence we have to choose a_{21}, b_1 and b_2 such that $c(h)-q = O(h^3)$.

Let us recall the classical order conditions for Runge–Kutta methods; see for example [3]. Let $A \in \mathbb{R}^{r\times r}$ be the matrix containing the coefficients of the (r–stage) method. Let $b = (b_1, \dots, b_r)$ and $\mathbf{1} = (1, \dots, 1)$. Let us denote the componentwise (or Schur or Hadamard) product of vectors by $a \diamond b = (a_1 b_1, \dots, a_r b_r)$. Then the order conditions for orders up to four are given by

$$\langle b, \mathbf{1}\rangle = 1, \tag{4.2}$$

$$\langle b, A\mathbf{1}\rangle = 1/2, \tag{4.3}$$

$$\begin{cases} \langle b, A^2\mathbf{1}\rangle = 1/6, \\ \langle b, A\mathbf{1} \diamond A\mathbf{1}\rangle = 1/3, \end{cases} \tag{4.4}$$

$$\begin{cases} \langle b, A^3\mathbf{1}\rangle = 1/24, \\ \langle b, A^2\mathbf{1} \diamond A\mathbf{1}\rangle = 1/8, \\ \langle b, A\mathbf{1} \diamond A\mathbf{1} \diamond A\mathbf{1}\rangle = 1/4, \\ \langle b, A(A\mathbf{1} \diamond A\mathbf{1})\rangle = 1/12. \end{cases} \tag{4.5}$$

PROPOSITION 4.1. *The scheme* (4.1) *is of order* 2 *if and only if the conditions* (4.2) *and* (4.3) *are satisfied.*

PROOF. Proceeding as in the previous section we readily get

$$\begin{aligned} \tilde q &= p + a_{21} V_p h + O(h^2), \\ V_{\tilde q} &= V_p + a_{21} dV\, V\, h + O(h^2). \end{aligned} \tag{4.6}$$

Then computing the expansion of q we first get

$$\begin{aligned} q^1 + (df_p)^t \mu^1 &= (b_1 + b_2) V_p, \\ df\, q^1 &= 0. \end{aligned}$$

Hence $\mu^1 = 0$ and $q^1 = (b_1 + b_2)V_p$ which implies $b_1 + b_2 = 1$, i.e. the condition (4.2). Proceeding further we get

$$\begin{aligned} q^2 + (df)^t \mu^2 &= b_2 a_{21} dV\, V, \\ df\, q^2 + \tfrac{1}{2}\, d^2 f(V, V) &= 0. \end{aligned}$$

For the second order method $q^2 = \frac{1}{2}\, dVV$. Hence $\mu^2 = 0$ and $b_2 a_{21} = 1/2$, i.e. condition (4.3) holds. □

Next let us consider a scheme with 3 stages.

$$\begin{cases} \tilde q + (df_{\tilde q})^t \tilde\mu = p + h a_{21} V_p, \\ f(\tilde q) = 0, \\ \hat q + (df_{\hat q})^t \hat\mu = p + h\big(a_{31} V_p + a_{32} V_{\tilde q}\big), \\ f(\hat q) = 0, \\ q + (df_q)^t \mu = p + h\big(b_1 V_p + b_2 V_{\tilde q} + b_3 V_{\hat q}\big), \\ f(q) = 0. \end{cases} \tag{4.7}$$

PROPOSITION 4.2. *The scheme* (4.7) *is of order* 3 *if and only if the conditions* (4.2), (4.3) *and* (4.4) *are satisfied.*

PROOF. We need more terms in the expansion (4.6). The second order terms give

$$\begin{aligned} &\tilde{q}^2 + (df)^t \tilde{\mu}^2 = 0, \\ &df\, \tilde{q}^2 + \tfrac{1}{2}\, a_{21}^2 d^2 f(V,V) = 0. \end{aligned}$$

Hence $B\tilde{\mu}^2 = -\frac{1}{2}\, a_{21}^2 df\, S_V$ and $\tilde{q}^2 = \frac{1}{2}\, a_{21}^2 S_V$. Next we must compute the expansion $\hat{q} = p + \hat{q}^1 h + \hat{q}^2 h^2 + O(h^3)$. Evidently $\hat{\mu}^1 = 0$ and $\hat{q}^1 = (a_{31} + a_{32}) V_p$ and for the second order terms we obtain

$$\begin{aligned} &\hat{q}^2 + (df_p)^t \hat{\mu}^2 = a_{32} a_{21} dV\, V, \\ &df\, \hat{q}^2 + \tfrac{1}{2}\, (a_{31} + a_{32})^2 d^2 f(V,V) = 0. \end{aligned}$$

This yields

$$\begin{aligned} B\hat{\mu}^2 =& df \Big(a_{32} a_{21} - \tfrac{1}{2}\, (a_{31} + a_{32})^2 \Big) S_V, \\ \hat{q}^2 =& a_{32} a_{21} \nabla_V V + \tfrac{1}{2}\, (a_{31} + a_{32})^2 S_V. \end{aligned} \tag{4.8}$$

Then we compute

$$\begin{aligned} V_{\tilde{q}} =& V_p + a_{21} dV\, V\, h + \tfrac{1}{2}\, a_{21}^2 \big([V, \nabla_V V] + dS_V V \big) h^2 + O(h^3), \\ V_{\hat{q}} =& V_p + (a_{31} + a_{32}) dV\, V\, h + \Big(a_{32} a_{21} \big(\nabla_V (\nabla_V V) + S_\nabla \big) \\ &+ \big(\tfrac{1}{2}\, (a_{31} + a_{32})^2 - a_{32} a_{21} \big) [V, \nabla_V V] \\ &+ \tfrac{1}{2}\, (a_{31} + a_{32})^2 dS_V V \Big) h^2 + O(h^3). \end{aligned} \tag{4.9}$$

Finally we must expand q. From the equations

$$\begin{aligned} &q^1 + (df_p)^t \mu^1 = (b_1 + b_2 + b_3) V, \\ &df\, q^1 = 0. \end{aligned}$$

we get $\mu^1 = 0$ and $q^1 = (b_1 + b_2 + b_3) V$ which gives the condition $b_1 + b_2 + b_3 = 1$. Expanding further we obtain

$$\begin{aligned} &q^2 + (df_p)^t \mu^2 = \big(b_2 a_{21} + b_3 (a_{31} + a_{32}) \big) dV\, V, \\ &df\, q^2 + \tfrac{1}{2}\, d^2 f(V,V) = 0. \end{aligned}$$

Requiring $q^2 = \frac{1}{2}\, dV\, V$ leads to $\mu^2 = 0$ and $b_2 a_{21} + b_3 (a_{31} + a_{32}) = 1/2$. Finally using (4.9) and (4.4) we get

$$\begin{aligned} q^3 + (df_p)^t \mu^3 =& \tfrac{1}{2} \left(b_2 a_{21}^2 + b_3 (a_{31} + a_{32})^2 \right) \big([V, \nabla_V V] + dS_V V \big) \\ &+ b_3 a_{32} a_{21} \big(\nabla_V (\nabla_V V) - [V, \nabla_V V] + S_\nabla \big), \\ =& \tfrac{1}{6} \Big(\nabla_V (\nabla_V V) + dS_V\, V + S_\nabla \Big). \end{aligned}$$

Comparing to Lemma 8.2 we conclude that $q^3 = \frac{1}{6}\, c'''(0)$ and $\mu^3 = 0$. □

Let us then consider a scheme with 4 stages.

$$
\begin{cases}
\tilde{q} + (df_{\tilde{q}})^t \tilde{\mu} = p + h a_{21} V_p, \\
f(\tilde{q}) = 0, \\
\hat{q} + (df_{\hat{q}})^t \hat{\mu} = p + h\big(a_{31} V_p + a_{32} V_{\tilde{q}}\big), \\
f(\hat{q}) = 0, \\
\bar{q} + (df_{\bar{q}})^t \bar{\mu} = p + h\big(a_{41} V_p + a_{42} V_{\tilde{q}} + a_{43} V_{\hat{q}}\big), \\
f(\bar{q}) = 0, \\
q + (df_q)^t \mu = p + h\big(b_1 V_p + b_2 V_{\tilde{q}} + b_3 V_{\hat{q}} + b_4 V_{\bar{q}}\big), \\
f(q) = 0.
\end{cases} \tag{4.10}
$$

PROPOSITION 4.3. *The scheme* (4.10) *is of order* 4 *if and only if the conditions* (4.2), (4.3), (4.4) *and* (4.5) *are satisfied.*

PROOF. We need more terms in various expansions. First we have to solve

$$
\begin{aligned}
&\tilde{q}^3 + (df)^t \tilde{\mu}^3 + a_{21} d^2 f(V, \cdot)\tilde{\mu}^2 = 0, \\
&df\, \tilde{q}^3 + \tfrac{1}{2}\, a_{21}^3 d^2 f(V, S_V) + \tfrac{1}{6}\, a_{21}^3 d^3 f(V, V, V) = 0.
\end{aligned}
$$

We obtain using Lemmas 8.1 and 8.4

$$
\begin{aligned}
B\tilde{\mu}^3 =& \tfrac{1}{6}\, a_{21}^3 d^3 f(V, V, V) + \tfrac{1}{2}\, a_{21}^3 d^2 f(V, S_V) + \tfrac{1}{2}\, a_{21}^3 df\, d^2 f(V, \cdot) B^{-1} df\, S_V, \\
=& \tfrac{1}{6}\, a_{21}^3 df\big(2S_\nabla - dS_V V - 3(dV)^t S_V\big), \\
\tilde{q}^3 =& \tfrac{1}{6}\, a_{21}^3 \big(3\pi_t(dS_V V) + \pi_n(dS_V V) - 2S_\nabla\big).
\end{aligned}
$$

Next we compute third order terms in the expansion of $\hat{q}$.

$$
\begin{aligned}
&\hat{q}^3 + (df_p)^t \hat{\mu}^3 + (a_{31} + a_{32}) d^2 f(V, \cdot)\hat{\mu}^2 = \tfrac{1}{2}\, a_{21}^2 a_{32}\big([V, \nabla_V V] + dS_V V\big), \\
&df\, \hat{q}^3 + \tfrac{1}{2}\, (a_{31} + a_{32}) d^2 f(V, \hat{q}^2) + \tfrac{1}{6}\, (a_{31} + a_{32})^3 d^3 f(V, V, V) = 0.
\end{aligned}
$$

$\hat{\mu}^2$ and $\hat{q}^2$ are given in (4.8). Solving this yields

$$
\begin{aligned}
B\hat{\mu}^3 =& df\Big(u_1 (dV)^t S_V + u_2 dS_V V + u_3 S_\nabla\Big), \\
\hat{q}^3 =& \tfrac{1}{2}\, a_{21}^2 a_{32}[V, \nabla_V V] + u_4 \pi_t(dS_V V) - u_3 S_\nabla + \tfrac{1}{6}\, (a_{31} + a_{32})^3 \pi_n(dS_V V),
\end{aligned}
$$

where u_i's are given by

$$
\begin{aligned}
u_1 =& a_{21} a_{32}(a_{31} + a_{32}) - \tfrac{1}{2}\, (a_{31} + a_{32})^3, \\
u_2 =& \tfrac{1}{2}\, a_{21}^2 a_{32} - \tfrac{1}{6}\, (a_{31} + a_{32})^3, \\
u_3 =& \tfrac{1}{3}\, (a_{31} + a_{32})^3 - a_{21} a_{32}(a_{31} + a_{32}), \\
u_4 =& a_{21} a_{32}(\tfrac{1}{2}\, a_{21} - a_{31} - a_{32}) + \tfrac{1}{2}\, (a_{31} + a_{32})^3.
\end{aligned}
$$

Then we move on to $\bar{q}$. Obviously $\bar{\mu}^1 = 0$ and $\bar{q}^1 = u_5 V_p$ where $u_5 = a_{41} + a_{42} + a_{43}$. The second order terms are solved from

$$
\begin{aligned}
&\bar{q}^2 + (df_p)^t \bar{\mu}^2 = \big(a_{42} a_{21} + a_{43}(a_{31} + a_{32})\big) dV\, V, \\
&df\, \bar{q}^2 + \tfrac{1}{2}\, u_5^2 d^2 f(V, V) = 0.
\end{aligned}
$$

Putting $u_6 = a_{42}a_{21} + a_{43}(a_{31} + a_{32})$ we get

$$\begin{aligned} B\bar{\mu}^2 =& df\Big(u_6 dV\, V - \tfrac{1}{2}\, u_5^2 S_V\Big), \\ \bar{q}^2 =& u_6 \nabla_V V + \tfrac{1}{2}\, u_5^2 S_V. \end{aligned}$$

Then the third order equations are

$$\begin{aligned} &\bar{q}^3 + (df_p)^t \bar{\mu}^3 + u_5 d^2 f(V,\cdot)\bar{\mu}^2 = u_7[V, \nabla_V V] + u_8 dS_V V \\ &\qquad + a_{21}a_{32}a_{43}\big(\nabla_V(\nabla_V V) + S_\nabla\big), \\ &df\, \bar{q}^3 + u_5 u_6 d^2 f(V, \nabla_V V) + \tfrac{1}{2}\, u_5^3 d^2 f(V, S_V) + \tfrac{1}{6}\, u_5^3 d^3 f(V,V,V) = 0, \end{aligned}$$

where

$$\begin{aligned} u_7 =& \tfrac{1}{2}\, a_{42}a_{21}^2 + \tfrac{1}{2}\, a_{43}(a_{31} + a_{32})^2 - a_{21}a_{32}a_{43}, \\ u_8 =& \tfrac{1}{2}\, a_{42}a_{21}^2 + \tfrac{1}{2}\, a_{43}(a_{31} + a_{32})^2. \end{aligned}$$

Then

$$\begin{aligned} B\bar{\mu}^3 =& df\Big((u_5u_6 - \tfrac{1}{2}\, u_5^3)(dV)^t S_V + (u_8 - \tfrac{1}{6}\, u_5^3) dS_V V \\ &+ (a_{21}a_{32}a_{43} + \tfrac{1}{3}\, u_5^3 - u_5u_6) S_\nabla\Big), \\ \bar{q}^3 =& u_7[V, \nabla_V V] + a_{21}a_{32}a_{43}\nabla_V(\nabla_V V) + (u_5u_6 - \tfrac{1}{3}\, u_5^3) S_\nabla \\ &+ (u_8 + \tfrac{1}{2}\, u_5^3 - u_5u_6)\pi_t(dS_V V) + \tfrac{1}{6}\, u_5^3 \pi_n(dS_V V). \end{aligned}$$

We shall need third order terms in the expansions (4.9) (denoted by $V^3_{\tilde{q}}$ and $V^3_{\hat{q}}$) as well as the expansion of $V_{\bar{q}}$.

$$\begin{aligned} V^3_{\tilde{q}} =& dV\, \tilde{q}^3 + \tfrac{1}{2}\, a_{21}^3 d^2V(V, S_V) + \tfrac{1}{6}\, a_{21}^3 d^3V(V,V,V), \\ V^3_{\hat{q}} =& dV\, \hat{q}^3 + a_{21}a_{32}(a_{31} + a_{32}) d^2V(V, \nabla_V V) \\ &+ \tfrac{1}{2}\, (a_{31} + a_{32})^3 d^2V(V, S_V) + \tfrac{1}{6}\, (a_{31} + a_{32})^3 d^3V(V,V,V), \end{aligned}$$

$$\begin{aligned} V_{\bar{q}} =& V_p + u_5 dV\, V\, h + \Big(u_6\big(\nabla_V(\nabla_V V) + S_\nabla\big) \\ &+ \big(\tfrac{1}{2}\, u_5^2 - u_6\big)[V, \nabla_V V] + \tfrac{1}{2}\, u_5^2 dS_V V\Big) h^2 \\ &+ \Big(dV\, \bar{q}^3 + u_5 u_6 d^2V(V, \nabla_V V) \\ &+ \tfrac{1}{2}\, u_5^3 d^2V(V, S_V) + \tfrac{1}{6}\, u_5^3 d^3V(V,V,V)\Big) h^3 + O(h^4). \end{aligned}$$

Finally we must expand q. First order terms are $\mu^1 = 0$ and $q^1 = (b_1 + b_2 + b_3 + b_4)V$, hence $\langle b, \mathbf{1}\rangle = 1$. Expanding further we obtain

$$\begin{aligned} &q^2 + (df_p)^t \mu^2 = \langle b, A\mathbf{1}\rangle dV\, V, \\ &df\, q^2 + \tfrac{1}{2}\, d^2 f(V,V) = 0. \end{aligned}$$

Requiring $q^2 = \frac{1}{2}\, dV\, V$ leads to $\mu^2 = 0$ and $\langle b, A\mathbf{1}\rangle = 1/2$. Then using (4.4) we get

$$\begin{aligned} q^3 + (df_p)^t\mu^3 =& \big(\tfrac{1}{2}\,\langle b, A\mathbf{1}\diamond A\mathbf{1}\rangle - \langle b, A^2\mathbf{1}\rangle\big)\,[V, \nabla_V V] \\ &+ \tfrac{1}{2}\,\langle b, A\mathbf{1}\diamond A\mathbf{1}\rangle dS_V V + \langle b, A^2\mathbf{1}\rangle\big(\nabla_V(\nabla_V V) + S_\nabla\big) \\ =& \tfrac{1}{6}\,\Big(\nabla_V(\nabla_V V) + dS_V\, V + S_\nabla\Big). \end{aligned} \tag{4.11}$$

Hence we conclude that $q^3 = \frac{1}{6}\, c'''(0)$ and $\mu^3 = 0$. Note that the fact that $\mu^3 = 0$ simplifies the computations in the last step. Consequently the equation for q^4 is

$$q^4 + (df)^t\mu^4 = b_2 V_{\tilde{q}}^3 + b_3 V_{\hat{q}}^3 + b_4 V_{\bar{q}}^3. \tag{4.12}$$

Anyway, at this point the computations became so tedious that it was necessary to use Mathematica [15]. Now guided by the previous steps we expect that $\mu^4 = 0$. Indeed, using the conditions (4.5) and simplifying we find that $b_2 V_{\tilde{q}}^3 + b_3 V_{\hat{q}}^3 + b_4 V_{\bar{q}}^3 = \frac{1}{24}\, c^{(4)}(0)$, where $c^{(4)}(0)$ is given in Lemma 8.7. □

4.2 Implicit methods.

The analysis follows the same lines as in the explicit case so we proceed here more rapidly. Let us start by the one-stage method

$$\begin{cases} \tilde{q} + (df_{\tilde{q}})^t\tilde{\mu} = p + ha_{11}V_{\tilde{q}}, \\ f(\tilde{q}) = 0, \\ q + (df_q)^t\mu = p + hb_1 V_{\tilde{q}}, \\ f(q) = 0. \end{cases} \tag{4.13}$$

PROPOSITION 4.4. *The scheme* (4.13) *is of order* 1 *if and only if* $b_1 = 1$ *(condition* (4.2)*) and of order* 2 *if and only if in addition* $a_{11} = 1/2$ *(condition* (4.3)*).*

PROOF. We easily get $\tilde{\mu} = O(h^2)$ and

$$\begin{aligned} \tilde{q} =& p + a_{11} V\, h + O(h^2), \\ V_{\tilde{q}} =& V_p + a_{11} dV\, V\, h + O(h^2). \end{aligned}$$

Hence $\mu^1 = 0$ and $q^1 = b_1 V_p = V_p$. The second order terms give

$$\begin{aligned} q^2 + (df_p)^t\mu^2 &= a_{11} dV\, V, \\ df\, q^2 + \tfrac{1}{2}\, d^2 f(V, V) &= 0. \end{aligned}$$

Hence the scheme is of second order if $a_{11} = 1/2$ in which case $\mu^2 = 0$. □

Note that the amount of work would be saved if $\tilde{q} = p$ ($a_{11} = 0$, explicit Euler) or $\tilde{q} = q$ ($a_{11} = b_1$, implicit Euler), but the maximal order is obtained only with $a_{11} = 1/2$ (midpoint rule).

Let us then consider the general two stage method.

$$
\begin{cases}
\tilde{q} + (df_{\tilde{q}})^t \tilde{\mu} = p + h\big(a_{11} V_{\tilde{q}} + a_{12} V_{\hat{q}}\big), \\
f(\tilde{q}) = 0, \\
\hat{q} + (df_{\hat{q}})^t \hat{\mu} = p + h\big(a_{21} V_{\tilde{q}} + a_{22} V_{\hat{q}}\big), \\
f(\hat{q}) = 0, \\
q + (df_q)^t \mu = p + h\big(b_1 V_{\tilde{q}} + b_2 V_{\hat{q}}\big), \\
f(q) = 0.
\end{cases} \tag{4.14}
$$

Proposition 4.5. *The scheme* (4.14) *is of order* 3 *if and only if conditions* (4.2), (4.3) *and* (4.4) *are satisfied.*

Proof. The proof is entirely analoguous to the explicit case; see [12] for details. □

Here again we can save the amount of work if we choose either $a_{11} = a_{12} = 0$ or $a_{21} = b_1$ and $a_{22} = b_2$. Here it is possible to have a third order method with these choices. In the first case we obtain $a_{21} = a_{22} = 1/3$, $b_1 = 1/4$ and $b_2 = 3/4$, and in the second case $a_{11} = 5/12$, $a_{12} = -1/12$, $a_{21} = 3/4$ and $a_{22} = 1/4$. The former scheme is known as RADAU I and the latter as RADAU IIA [3].

Let us then consider the general three stage method.

$$
\begin{cases}
\tilde{q} + (df_{\tilde{q}})^t \tilde{\mu} = p + h\big(a_{11} V_{\tilde{q}} + a_{12} V_{\hat{q}} + a_{13} V_{\bar{q}}\big), \\
f(\tilde{q}) = 0, \\
\hat{q} + (df_{\hat{q}})^t \hat{\mu} = p + h\big(a_{21} V_{\tilde{q}} + a_{22} V_{\hat{q}} + a_{23} V_{\bar{q}}\big), \\
f(\hat{q}) = 0, \\
\bar{q} + (df_{\bar{q}})^t \bar{\mu} = p + h\big(a_{31} V_{\tilde{q}} + a_{32} V_{\hat{q}} + a_{33} V_{\bar{q}}\big), \\
f(\bar{q}) = 0, \\
q + (df_q)^t \mu = p + h\big(b_1 V_{\tilde{q}} + b_2 V_{\hat{q}} + b_3 V_{\bar{q}}\big), \\
f(q) = 0.
\end{cases} \tag{4.15}
$$

Proposition 4.6. *The scheme* (4.15) *is of order* 4 *if and only if conditions* (4.2), (4.3), (4.4) *and* (4.5) *are satisfied.*

Proof. See [12]. □

Again it is possible to save the amount of work and to have order four by RADAU type methods, i.e. by requiring that either $a_{1j} = 0$ or $a_{3j} = b_j$. In fact one can impose both conditions and keep order four; these kind of schemes are known as LOBATTO IIIA. So even though in this case there are 8 conditions and only 6 parameters, there happens to be a unique solution: $a_{21} = 5/24$, $a_{22} = 1/3$, $a_{23} = -1/24$, $a_{31} = 1/6$, $a_{32} = 2/3$, $a_{33} = 1/6$.

To sum up, all the results in this section suggest that the following conjecture could be true.

Conjecture 4.7. *Our versions of Runge–Kutta methods have the same order as the classical versions.*

5 Numerical implementation.

Let us consider the system (2.6). Given an initial point $p \in \mathcal{R}_q = f^{-1}(0) \subset J_q(\mathcal{E})$ we would like to compute the integral manifold going through p. There are three subproblems in implementing the algorithm:

1. Given a point $p \in \mathcal{R}_q$, compute the distribution $\mathcal{D}_p$ which is defined by (2.5).
2. Step size control.
3. When the (intermediate) steps give points which are outside $\mathcal{R}_q$ they must be orthogonally projected back to $\mathcal{R}_q$.

All of these reduce to fairly standard numerical problems. The implementation was done on Maple V [4]. It would be more efficient to produce, e.g., Fortran code from the Maple procedures, but this will be considered in a future paper.

5.1 Computing the distribution.

At each point the distribution can be represented as a nullspace of the following matrix [13]:

$$A = \begin{pmatrix} -v & I_{nq} & 0_{nq\times n} \\ w & A_1 & A_2 \end{pmatrix},$$

where $w = \partial f/\partial x$, $A_2 = \partial f/\partial y_q$, A_1 contains the partial derivatives of f with respect to $y, y_1, \ldots, y_{q-1}$ and finally $v \in \mathbb{R}^{nq\times 1}$ contains the vectors y_1, y_2, $\ldots$, y_q. The computation of the nullspace of A can further be reduced to the computation of the nullspace of the following $k \times (n+1)$-matrix:

$$C = \begin{pmatrix} w + A_1 v & A_2 \end{pmatrix}. \tag{5.1}$$

Note that the dimensions of C are independent of q. For the computation of the nullspace we used the standard singular value decomposition $C = U\Sigma V^t$. The last singular value should be very close to zero, and hence the last column of V gives a good approximation of the required direction.

Note that for an important class of systems, namely mechanical systems with holonomic constraints, the distribution can be computed from the regular linear system and the SVD is not needed at all; see [13].

5.2 Step size control.

Here we have used two standard techniques [7]: either take two sequential steps with stepsize h and compare the result with one step with stepsize $2h$, or Fehlberg-like: compare two parallel steps, the other of order p and the other of order $p+1$.

We implemented the following explicit methods: classical RK4, Fehlberg 4(5) (denoted by RKF45) and Dormand–Prince 5(4) (denoted by dopri54). In case of RK4 the '$2 * h$ compared to $2h$' was used.

The number of steps that were rejected was relatively high, typically 10–20%. Perhaps the standard step selection strategies in [7] should be modified somewhat in the present context. This problem will be addressed in a future work.

5.3 Projection.

Projection was done by chord Newton iteration. The needed Jacobian was evaluated every third step and the resulting linear system was solved using LU-decomposition. Initial point could be selected either as

$$\begin{cases} q_{init} = p + hV_p, \\ \mu_{init} = 0, \end{cases} \tag{5.2}$$

or yet better as

$$\begin{cases} q_{init} = p + hV_p - (df_p)^t a\, h^2, \\ \mu_{init} = a\, h^2 = (df_p(df_p)^t)^{-1}\, d^2 f_p(V_p, V_p) h^2/2. \end{cases} \tag{5.3}$$

The distance from the initial point in (5.2) to the correct value is of order $O(h^2)$ and in (5.3) of order $O(h^3)$. In the numerical examples we use the initial point (5.3). The iteration was continued until the difference of successive iterates in maximum norm was smaller than some given tolerance, typically about 10^{-7}, or until the maximum number of iterations (30 was used) was exceeded. In the latter case step-size was reduced and then the computation was resumed. However, the non-convergence was extremely rare and usually only about 3 iterations was required, hence normally the Jacobian of f was evaluated only once in each projection. When the initial point was as in (5.2), typically one more iteration was needed, and so it seems to be advantageous to use the more complicated initial point (5.3). If the tolerances in the step-size control were large, then the number of iterations naturally tended to increase.

REMARK 5.1. The h^2-coefficient of q_{init} in (5.3) has the following geometrical interpretation:

$$|(df_p)^t a| = \frac{1}{2} \cdot (\text{normal curvature at } p).$$

REMARK 5.2. When computing the distribution we need the Jacobian of f and in the Newton iteration as well as in the computation of the initial guess (5.3) we need in addition the second differential $d^2 f$. These were computed symbolically and then evaluated when needed. In the moderate sized examples below this was not a problem, but in bigger problems these computations and evaluations are eventually quite time consuming. Fortunately there is an efficient way to compute the required quantities using automatic differentiation [2]. However, a discussion of this technique is outside the scope of the present article and will be considered elsewhere.

6 Examples.

6.1 A simple example.

Let us look at a simple system in $J_1(\mathbb{R} \times \mathbb{R})$: $y_1 - 3y - 2x^2 = 0$, whose explicit solution with initial point $y(0) = 2$ is

$$y(x) = -x^2 - \frac{2}{3}x - \frac{2}{9} + \frac{20}{9}e^{3x}.$$

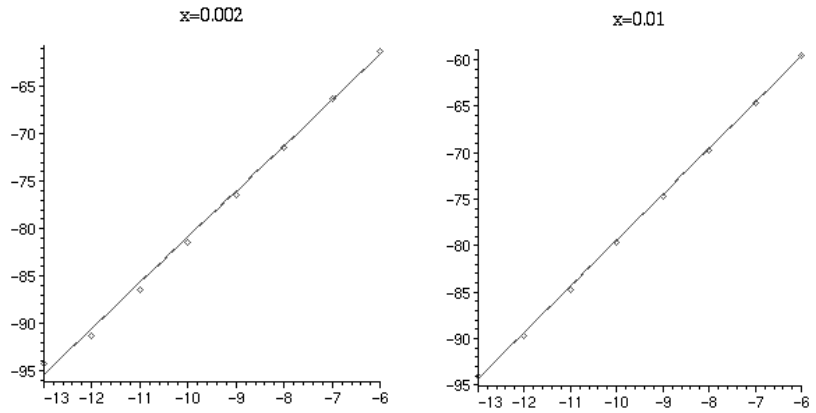

Figure 6.1: Testing the order of dopri54.

The motivation for this example is to test the order of the method dopri54. Note that the theory is so far only up to order 4 and we would like to test numerically the validity of Conjecture 4.7. We computed the solution for $x = 0, \ldots, 0.01$ with constant stepsize $h = 2^{-k}$ where $k \in \{6, 7, \ldots, 13\}$.

The result is encouraging, the order really seems to be 5. In Figure 6.1 the errors are shown in log-log scale at $x = 0.002$ and $x = 0.01$. The slopes are 4.836 and 4.952, respectively. At $h = 2^{-13}$ one can see the effect of round-off errors, the computations were done with 30 decimals. Without the points where $h = 2^{-13}$ the slopes would be 5.008 and 5.011, respectively. So there is also some numerical evidence to support Conjecture 4.7.

6.2 Hénon–Heiles system.

The equations of this famous example from astronomy are:

$$\begin{cases} y_2^1 + y^1 + 2y^1y^2 = 0, \\ y_2^2 + y^2 + (y^1)^2 - (y^2)^2 = 0, \\ \frac{1}{2}\,|y_1|^2 + \frac{1}{2}\,|y|^2 + (y^1)^2y^2 - \frac{1}{3}(y^2)^3 - a = 0. \end{cases} \tag{6.1}$$

The computation is reduced [13] to:

$$\begin{cases} \frac{1}{2}\,|y_1|^2 + \frac{1}{2}\,|y|^2 + (y^1)^2y^2 - \frac{1}{3}(y^2)^3 - a = 0, \\ \mathcal{D} = \operatorname{span}(V), \\ V = \big(1, y_1^1, y_1^2, -y^1 - 2y^1y^2, -y^2 - (y^1)^2 + (y^2)^2\big), \end{cases} \tag{6.2}$$

where a is the constant energy. We are interested in quasiperiodic motion, which is achieved by the initial point $(x, y, y_1) = (0, 0.12, 0.12, 0.12, 0.12)$. We take the Poincaré section at the hyperplane $y^1 = 0$ and look at $(y^2,\, y_1^2)$.

Table 6.1: Some results of Hénon–Heiles (6.2): number of steps.

	tol	x=110	x=550	x=1100
dopri54 4-th order pt	$5{\cdot}10^{-5}$	186	920	1837
dopri54 5-th order pt	$1{\cdot}10^{-5}$	242	1201	2404
dopri54 4-th order pt	$5{\cdot}10^{-6}$	268	1373	2722
RK4	$5{\cdot}10^{-4}$	568	2897	5801

It is well known [10] that in case of Hamiltonian systems the symplectic Runge–Kutta methods are better than ordinary Runge–Kutta methods, especially if the simulation times are long. However, as seen below, our method, without being symplectic, also produces qualitatively correct results on long time intervals.

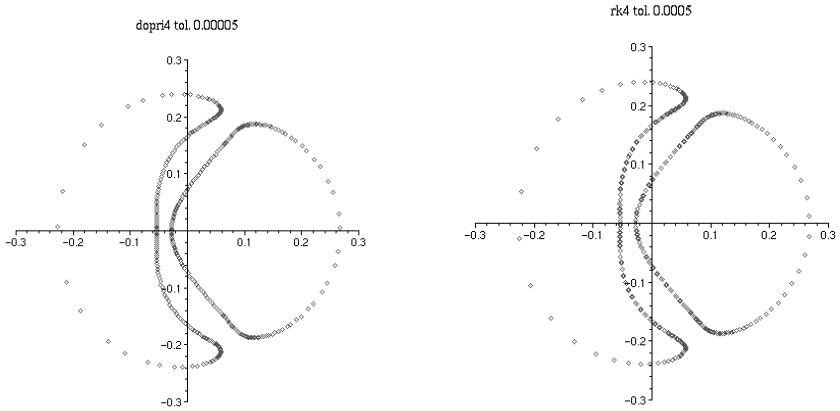

Figure 6.2: Poincaré sections in the (y^2, y_1^2)-plane for the Hénon–Heiles system (6.2).

We compute 'approximately two rounds' in the Poincaré section, which means taking $x = 0, \ldots, 1100$ and hence about 340 points in the section. The point in the section is chosen by Hermite interpolation between points p_m and p_{m+1} which are sequential such that the sign of y^1 changes.

With dopri54 we used the 4-th order point as the new starting point for the next step. Although it was not extensively studied, it seemed like using 5-th order point required so small a tolerance that proceeding along the solution was actually slower than in the 4-th order case.

Dopri54 was clearly more efficient than RKF45. Surprisingly, also RK4 was clearly better than RKF45, despite the more elementary step size control. RKF45 required very small stepsize to produce qualitatively correct solution, so it was not reasonable to compute even one round. Compared to this dopri54 and RK4 used quite big steps. The qualitative performance of RKF45 did not depend on choosing 4-th or 5-th order value for the continuation of the solution. Some results of computations are in Table 6.1.

On the left of Figure 6.2 is the result with dopri54, tol= $5\cdot 10^{-5}$, for which 1840 points was needed. With bigger tolerances the algorithm rejected considerably more points. By looking at the figure, the result is slightly less 'uniformly distributed' than in the following RK4 case:

On the right of Figure 6.2 is the result with RK4, tol= $5\cdot 10^{-4}$, for which 5800 points was needed. It is interesting to compare these two figures: the points in the sections are clearly not at same places, but the shapes of the sections are still the same.

REMARK 6.1. Note that we have not used any ideas from the symplectic geometry. This is because the relevant flows restricted to our manifolds would not be symplectic anymore, and hence it is not clear how to exploit the 'remaining' symplectic structure in the computations. However, it is quite possible that a deeper study of the connections of the symplectic and jet geometry would turn out to be interesting also from the numerical point of view.

6.3 Stiff pendulum.

This is a pendulum with massless, stiff spring with a spring constant $1/\varepsilon^2$, $\varepsilon > 0$ and a mass of 1 unit at the end of the spring. The rest length of the spring and the gravitational constant are taken to be 1. The equations are:

$$
\begin{cases}
y^1_2 + y^1 y^3 = 0,\\
y^2_2 + y^2 y^3 + 1 = 0,\\
\big((y^1)^2 + (y^2)^2\big)\big(\varepsilon^2 y^3 - 1\big)^2 - 1 = 0,\\
\varepsilon^2 y^3_1 + \big(y^1 y^1_1 + y^2 y^2_1\big)\big(\varepsilon^2 y^3 - 1\big)^3 = 0,\\
\big(\varepsilon^2 y^3 - 1\big)\varepsilon^2 y^3_2 - 3\varepsilon^4 (y^3_1)^2,\\
\quad + \big((y^1_1)^2 + (y^2_1)^2 - y^2\big)\big(\varepsilon^2 y^3 - 1\big)^4 - y^3\big(\varepsilon^2 y^3 - 1\big)^2 = 0,
\end{cases}
\tag{6.3}
$$

and the computation in reduced form [13]:

$$
\begin{cases}
\big((y^1)^2 + (y^2)^2\big)\big(\varepsilon^2 y^3 - 1\big)^2 - 1 = 0,\\
\mathcal{D} = \operatorname{span}(V),\\
V = \big(1, y^1_1, y^2_1, -\big(y^1 y^1_1 + y^2 y^2_1\big)\big(\varepsilon^2 y^3 - 1\big)^3/\varepsilon^2, -y^1 y^3, -y^2 y^3 - 1\big).
\end{cases}
\tag{6.4}
$$

Table 6.2: Stiff pendulum (6.4): $\varepsilon = 0.01$, number of steps.

	toler	$x = 0.5$	$x = 1.5$
dopri54	10^{-4}	357	1141
dopri54	10^{-5}	685	2051
RK4	10^{-2}	643	1931
RK4	10^{-3}	957	2857

We investigate the cases $\varepsilon \in \{0.1, 0.01\}$. The ε is a stiffness parameter: the smaller the epsilon, the stiffer the equations.

Let us choose first $\varepsilon = 0.01$ and use the initial point $(0, 0.85, 0, -1765, 0, 0)$. Note that the initial point does not have to satisfy the equations to high accuracy. We compute the solution in the interval $x = [0, 1.5]$.

Dopri54 is, as in previous example, clearly more efficient than RKF45. The latter has trouble keeping the solution qualitatively correct, it suffers from dissipativity (Figure 6.3) for all reasonable stepsizes. Some results of computations are in Table 6.2.

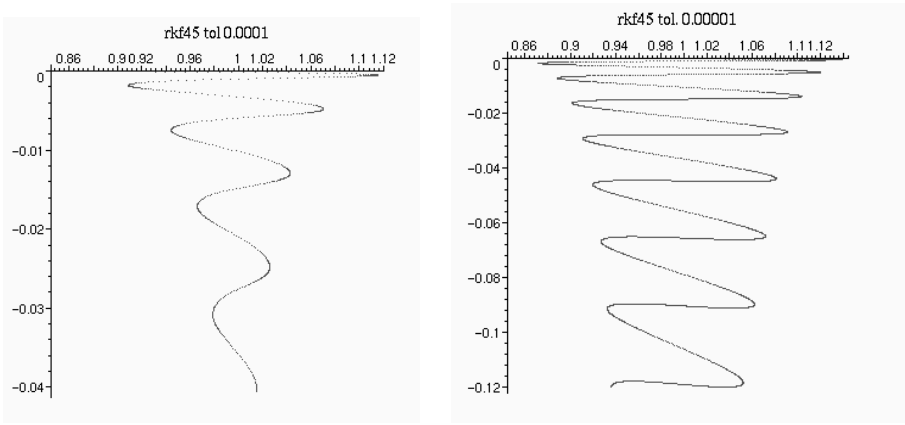

Figure 6.3: Dissipativity of RKF45 in stiff pendulum (6.4), $\varepsilon = 0.01$.

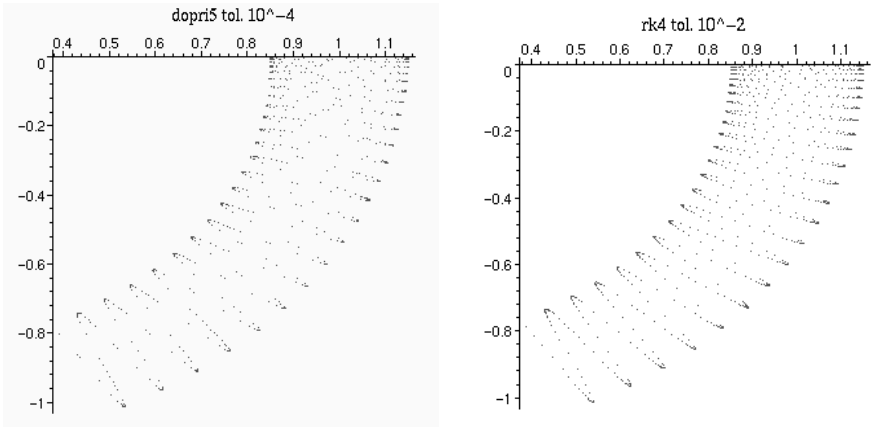

Figure 6.4: Stiff pendulum (6.4) $\varepsilon = 0.01$, $x = 0, \ldots, 1.5$.

Comparing dopri54 and RK4: on the left of Figure 6.4 is the result with dopri54

tol $= 10^{-4}$, we use the usual 5-th order point as the new point on solution. 1141 points was needed. With tolerances bigger than that, dopri54 lost the qualitatively correct behaviour. With tolerances 10^{-6} or smaller, the propagation along the solution was painfully slow.

Table 6.3: Stiff pendulum (6.4): $\varepsilon = 0.1$, number of steps.

	tol	$x = 3.7$	$x = 20$
dopri54	10^{-4}	136	751
dopri54	10^{-5}	260	1444
RK4	10^{-2}	223	1259
RK4	10^{-3}	371	2082

In RK4 the computed solution is of good quality even with 'big' tolerances. Hence in this example RK4 easily beats RKF45. RK4 is also more reliable but slower than dopri54. On the right of Figure 6.4 is the result with RK4 with tolerance equal to 10^{-2}. In dopri54 the effect of tolerance is clearly more visible than in RK4 case.

Let us choose then $\varepsilon = 0.1$, the initial point $p_0 = (0, .85, 0, -17.65, 0, 0)$ and compute the solution on the interval $x = [0, 20]$. The results are quite similar to the previous case; RK4 beats RKF45 and dopri54 is clearly superior to both of these. RKF45 is dissipative for all reasonable stepsizes.

Some results of computations are in Table 6.3. The column $x = 3.7$ corresponds to the 'one swing' case mentioned below.

In Figure 6.5 is the result by dopri54 with tol$= 10^{-4}$, on the left only one swing to show the 'speed' of numerical solution, on the right the qualitative behaviour over longer time interval.

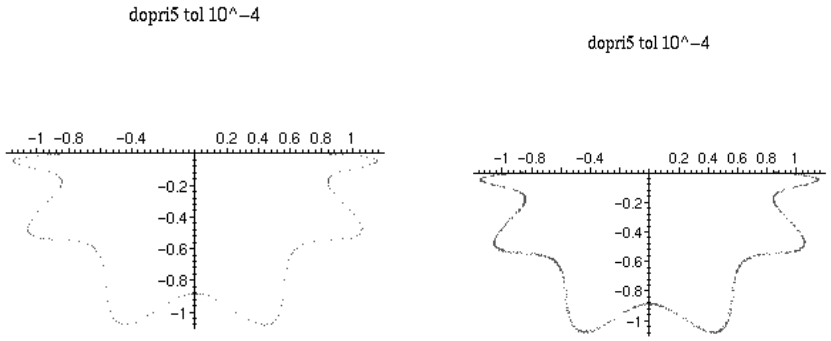

Figure 6.5: Stiff pendulum (6.4), $\varepsilon = 0.1$.

Remark 6.2. The surprisingly bad performance of RKF45 in this and the previous example is somewhat mysterious: we did not find any theoretical explanation for this fact.

6.4 Four bar system.

This is a classical example in multibody dynamics; see Figure 6.6. The links are rigid and of negligible mass. Joints are frictionless and joints 0 and 3 are fixed. There are point masses m_1, m_2 at joints 1 and 2, respectively. The effecting forces are gravity and a constant torque T acting on the origin.

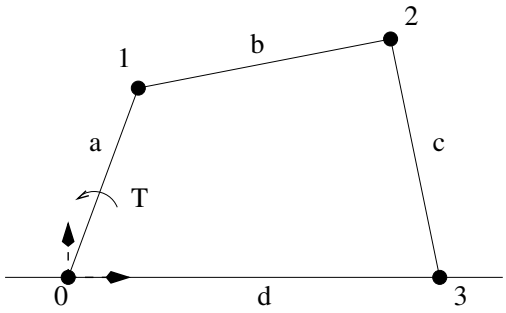

Figure 6.6: Four bar linkage.

The equations of motion, where (y^1, y^2), (y^3, y^4) are coordinates for joints 1 and 2 respectively, are in descriptor form:

$$\begin{cases} B(x, y, y_1)y_2 + f(x, y, y_1) + (dg)^t\lambda = 0, \\ g(y) = 0, \end{cases} \tag{6.5}$$

with $B = \mathrm{diag}(m_1, m_1, m_2, m_2)$, $\gamma =$ gravitation constant and

$$f = \begin{pmatrix} y^2\, T/a^2 \\ -y^1\, T/a^2 + m_1\gamma \\ 0 \\ m_2\gamma \end{pmatrix}, \quad g = \frac{1}{2}\begin{pmatrix} (y^1)^2 + (y^2)^2 - a^2 \\ (y^3 - y^1)^2 + (y^4 - y^2)^2 - b^2 \\ (d - y^3)^2 + (y^4)^2 - c^2 \end{pmatrix}. \tag{6.6}$$

In addition to these, the energy of the system is constant and expressed by

$$E = \frac{1}{2}\left(m_1((y_1^1)^2 + (y_1^2)^2) + m_2((y_1^3)^2 + (y_1^4)^2)\right) + (m_1 y^2 + m_2 y^4)\gamma \tag{6.7}$$

The constraints are holonomic, hence we can use in computations the reduced form as explained in [13]:

$$\begin{cases} g(y) = 0, \\ dg\, y_1 = 0, \\ \mathcal{D} = \mathrm{span}(V), \\ V = \big(1, y_1, y_2\big), \\ \begin{pmatrix} B & (dg)^t, \\ dg & 0 \end{pmatrix}\begin{pmatrix} y_2, \\ \lambda \end{pmatrix} + \begin{pmatrix} f, \\ d^2g\,(y_1, y_1) \end{pmatrix} = 0. \end{cases} \tag{6.8}$$

We shall consider the motion of joint 2 with different choices of a, b, c, d. Note that augmenting the system by constant energy condition changes the qualitative behaviour of the solution quite radically as expected since the variations in the energy are quite big. Let us recall that in the multibody system in [13] we had an example where a solution remained quite close to the energy surface although it was completely different from the solution obtained by imposing the constant energy condition.

Choose $(a, b, c, d) = (1, 2, 1.5, 2)$ and $T = 0.8$, $m_1 = 0.2$, $m_2 = 0.1$. We look at component y^4. The initial point is $(0, 0.5, 0.866, 2.4155, 1.4413\,, 0, 0, 0, 0)$. In Figure 6.7 is the result with the energy equation augmented (solid line) and without it (dashed line). The change of energy in the latter case is shown in figure 6.8.

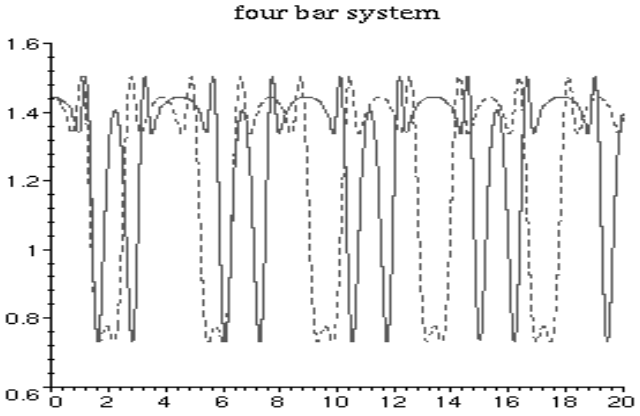

Figure 6.7: Results of four bar system (6.6) without (dashed line) and with (solid line) (6.7), $T = 0.8$.

In the constant energy case the result is beautifully oscillatory, as one might heuristically expect, since the effecting forces (gravitation and torque T) are constants. In the non-constant energy case the result is likewise oscillatory but the period of the oscillation is clearly shorter than in the 'correct' case of constant energy. Also, in down position the behaviour is quite different: it makes only a small 'cup' there.

The case where $a + b = c + d$ is interesting, since it produces singularities: then the linkage is capable of reaching position where all the bars are collinear. In such a case, the rank of dg is not maximal and the equations of motion become singular. This is called *constraint singularity*. The physical interpretation for the singularity is that there is a bifurcation: the system can move either both joints 1 and 2 down (or up) or another joint moves upwards while the other one moves downwards.

However, if we are modelling a real mechanical device, we might expect that it has some (non-modelled) supportive structures that eliminate the possibility of a bifurcation. In this case, we have to modify the equations to remove the

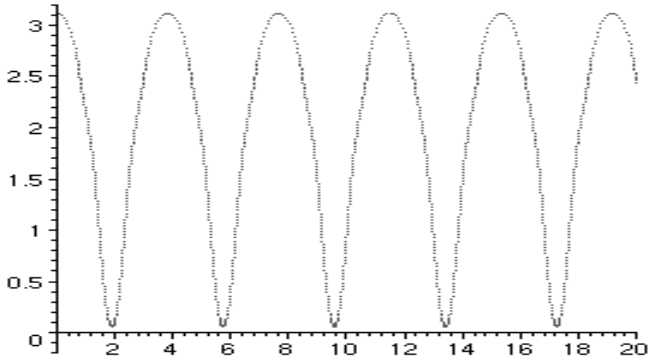

Figure 6.8: Evolution of energy of four bar system (6.6), $T = 0.8$.

singularity also from the equations. See also [1, 8] for other approaches to this problem.

The technique for resolving this problem is *ideal decomposition*. We need some definitions and results (which we state without proofs) from algebraic geometry [5]:

DEFINITION 6.1. *Denote by* $\mathbb{C}[y_1, \ldots, y_n]$ *the ring of polynomials in* $y_1, \ldots, y_n$. *A subset* $\mathcal{I} \subset \mathbb{C}[y_1, \ldots, y_n]$ *is an* ideal *if it satisfies*

(i) $0 \in \mathcal{I}$.

(ii) *If* $f, g \in \mathcal{I}$, *then* $f + g \in \mathcal{I}$.

(iii) *If* $f \in \mathcal{I}$ *and* $h \in \mathbb{C}[y_1, \ldots, y_n]$, *then* $hf \in \mathcal{I}$.

DEFINITION 6.2. *Let* $f_1, \ldots, f_s \in \mathbb{C}[y_1, \ldots, y_n]$. *Then we set*

$$\langle f_1, \ldots, f_s \rangle = \Big\{ \sum_{i=1}^{s} h_i f_i : h_1, \ldots, h_s \in \mathbb{C}[y_1, \ldots, y_n] \Big\}.$$

We call $\langle f_1, \ldots, f_s \rangle$ *the* ideal generated by $f_1, \ldots, f_s$.

We skip the proof of the fact that $\langle f_1, \ldots, f_s \rangle$ really is an ideal.

DEFINITION 6.3. *An ideal* $\mathcal{I}$ *is* radical *if* $f^m \in \mathcal{I}$ *for any integer* $m \geq 1$ *implies that* $f \in \mathcal{I}$. *An ideal* $\mathcal{I}$ *is* prime *if whenever* $f, g \in \mathbb{C}[y_1, \ldots, y_n]$ *and* $fg \in \mathcal{I}$, *then either* $f \in \mathcal{I}$ *or* $g \in \mathcal{I}$.

DEFINITION 6.4. *Let* $\mathcal{I}$ *be an ideal. The* radical *of* $\mathcal{I}$, *denoted by* $\sqrt{\mathcal{I}}$, *is the set*

$$\{f \in \mathbb{C}[y_1, \ldots, y_n] : f^m \in \mathcal{I} \text{ for some integer } m \geq 1\}.$$

LEMMA 6.1. *If $\mathcal{I}$ is an ideal, then $\sqrt{\mathcal{I}}$ is a radical ideal.*

Finally, a strong theorem from algebraic geometry:

THEOREM 6.2. *Every radical ideal $\mathcal{I}$ in $\mathbb{C}[y_1, \ldots, y_n]$ can be written uniquely as a finite intersection of prime ideals,*

$$\mathcal{I} = \mathcal{I}_1 \cap \cdots \cap \mathcal{I}_r,$$

where $\mathcal{I}_i \not\subset \mathcal{I}_j$ for $i \neq j$.

In our application there is no harm in restricting our attention to radical ideals, because we are interested in the zero sets defined by the generators and obviously the zero set defined by an ideal is the same as the zero set of its radical. An essential thing is that this decomposition can be computed algorithmically through the use of the generators of the ideal.

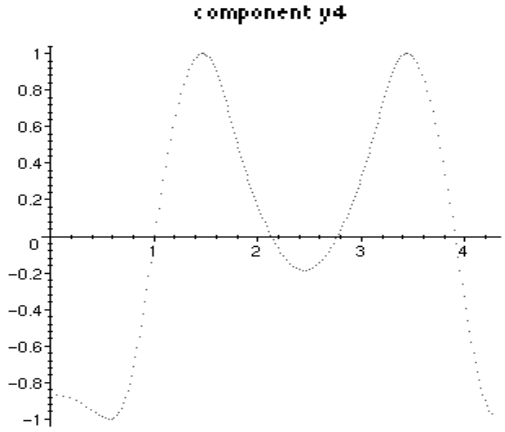

Figure 6.9: The component y^4 of the 'singular' four bar linkage.

To demonstrate, suppose $(a, b, c, d) = (1, 2, 1, 2)$. Our g is a polynomial, hence it generates an ideal $\mathcal{I} \in \mathbb{C}[y^1, y^2, y^3, y^4]$ which then generates a *radical* ideal $\sqrt{\mathcal{I}}$. This decomposes to intersection of prime ideals:

$$\sqrt{\mathcal{I}} = \mathcal{I}_1 \cap \mathcal{I}_2$$

where $\mathcal{I}_i$ is the prime ideal generated by g_i,

$$g_1 := \begin{cases} (y^3)^2 + (y^4)^2 - 4y^3 + 3, \\ y^2 - y^4, \\ y^1 - y^3 + 2, \end{cases}$$

$$g_2 := \begin{cases} (y^3)^2 + (y^4)^2 - 4y^3 + 3, \\ 4y^2y^3 - 3y^2 + 3y^4, \\ 4y^2y^4 + 3y^1 - 3y^3 + 6. \end{cases}$$

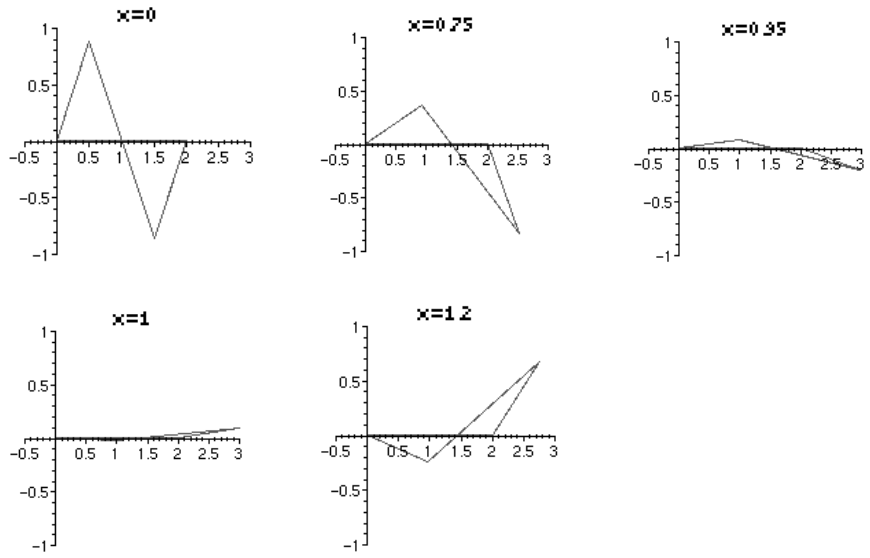

Figure 6.10: Some configurations of the 'singular' four bar linkage.

In computing these we found Singular [6] very helpful. Using these equations actually removes the singularity! For example, if we have the parallelogram case where both joints 1 and 2 move at the same height, we use g_1 in place of g. The situation $y^2 = y^4 = 0$ is no longer a singularity, g_1 is of maximal rank everywhere. Of course, in this simple case we can eliminate y^3 and y^4, a geometric fact which is also clearly visible in g_1.

Let's look more closely at g_2. The condition $y^2 = y^4$ implies $y^2 = y^4 = 0$. Hence the parallelogram motion is not possible in this case. On the other hand the middle equation of g_2 gives $y^4 = y^2(1 - \frac{4}{3}y^3)$ where the multiplier of y^2 is negative because the first equation implies that $1 \leq y^3 \leq 3$. That is, when nonzero, y^2 and y^4 have different signs and so g_2 corresponds to situations where the joints 1 and 2 are at different sides of the d-bar, as expected.

Note that the rank of g_2 is maximal everywhere and hence the singularity is vaporized! Let us take an example using g_2 in place of g: the initial point is

$$(0,\, 0.5,\, \sqrt{3}/2,\, 1.5,\, -\sqrt{3}/2, 0, 0, 0, 0)$$

and T, m_1, m_2 are as before. The component y^4 is in Figure 6.9 and some of the configurations of the linkage are in Figure 6.10. We use short stepsize in computation to assure that we get close to points where $y^4 = 0$. There is no problem in passing through points $y^4 = 0$ which would be singularities when using the original g.

REMARK 6.3. If we had used the Lagrange 2nd kind of equations (i.e. using a minimal set of coordinates, e.g. an angle between bars a and d) for representing the equations of motion, we would have less equations but they would still include

an algebraic constraint. Moreover, this constraint would be of non-polynomial type and the ideal decomposition could not be applied.

7 Conclusion.

We have shown that it is possible to implement explicit higher order Runge–Kutta type schemes for computing solutions of overdetermined systems or DAEs. As explained in more detail in [13] we do not encounter the numerical problems usually associated to these problems. The difficulties in efficient implementation are of different nature. The computation of the involutive form of the system which is needed may be quite time consuming because usually some symbolic computation is required. On the other hand for quite large and interesting class of systems, namely mechanical systems with holonomic constraints and Hamiltonian systems, the involutive form is easily available. It is an interesting open problem to study how to exploit specific structures of systems in the computation of the involutive form, and if some kind of mixture of symbolic and numerical computation would be sufficient, at least in certain cases, which could result in significant speed-ups.

On the more numerical side, the implementation of the subproblems in our code, namely computation of the distribution and projection, are not as efficient as they could be. A careful study of these problems will be a subject of future papers. Anyway, and perhaps most importantly, our computations show that solutions to overdetermined systems can be computed to high accuracy in a numerically stable way; in fact the overdetermined character of the system is barely visible in our formulation.

8 Auxiliary lemmas.

Let $f : \mathbb{R}^m \mapsto \mathbb{R}^k$, $B = df(df)^t$ and $M := f^{-1}(0) \subset \mathbb{R}^m$ as before. V will always be a vector field on M, i.e. $V \in \Gamma(TM)$.

LEMMA 8.1. *Let $Y \in \Gamma(T\mathbb{R}^m)$; then*

$$d^2 f(V, \cdot) = -(dV)^t (df)^t,$$
$$d^2 f(V, Y) = -df\, dV\, Y.$$

PROOF. We observe that since $\langle df^j, V\rangle = 0$

$$\left(d^2 f(V, \cdot)\right)_{i,j} = \sum_k \frac{\partial^2 f^j}{\partial x^i \partial x^k} v^k = \frac{\partial}{\partial x^i} \sum_k \frac{\partial f^j}{\partial x^k} v^k - \sum_k \frac{\partial f^j}{\partial x^k} \frac{\partial v^k}{\partial x^i}$$
$$= -\sum_k \frac{\partial f^j}{\partial x^k} \frac{\partial v^k}{\partial x^i}.$$

The proof of the other statement is similar. □

LEMMA 8.2. *Let $c : \mathbb{R} \mapsto M$ and $c(0) = p$; then*

$$c'''(0) = \nabla_V(\nabla_V V) + dS_V\, V + S_\nabla.$$

PROOF. Recall that $c''(0) = dVV = \nabla_V V + S_V$. Then we get the result by noting that $c'''(0) = d(\nabla_V V + S_V)V$ and using (2.3). □

LEMMA 8.3.

$$\pi_t(dS_V V) = -\pi_t((dV)^t S_V).$$

PROOF. Let $Y \in \Gamma(TM)$; then using (2.2) and (2.3) we get

$$0 = V_p\langle S_V, Y\rangle = \langle dS_V V, Y\rangle + \langle S_V, dYV\rangle = \langle dS_V V, Y\rangle + \langle S_V, S(Y,V)\rangle.$$

On the other hand

$$\langle (dV)^t S_V, Y\rangle = \langle S_V, dVY\rangle = \langle S_V, S(Y,V)\rangle.$$

□

LEMMA 8.4.

$$d^3 f(V,V,V) = df\big(3dV\, S_V - dS_V\, V + 2S_\nabla\big).$$

PROOF.

$$\begin{aligned}\big(d^3 f(V,V,V)\big)_i &= \sum_{j,k,l} \frac{\partial^3 f^i}{\partial x^j \partial x^k \partial x^l} v^j v^k v^l \\ &= \sum_l v^l \frac{\partial}{\partial x^l} \sum_{j,k} \frac{\partial^2 f^i}{\partial x^j \partial x^k} v^j v^k - 2\sum_{j,k,l} \frac{\partial^2 f^i}{\partial x^j \partial x^k} \frac{\partial v^j}{\partial x^l} v^k v^l \\ &= -\, d\big(df^i\, S_V\big)V - 2d^2 f(dV\, V, V).\end{aligned}$$

Hence

$$d^3 f(V,V,V) = -df\, dS_V\, V + 2df\, S_\nabla - 3d^2 f(S_V, V),$$

which combined with Lemma 8.1 gives the result. □

LEMMA 8.5. *Let $Y \in \Gamma(TM)$; then*

$$\begin{aligned}d^2V(V,Y) =& \nabla_Y(\nabla_V V) - \nabla_V(\nabla_Y V) + [V, \nabla_V Y] - [V, [V,Y]\,] \\ &+ dS_V Y - dV\, S(V,Y) + S(Y, \nabla_V V) - S(V, \nabla_Y V).\end{aligned}$$

In particular $d^2V(V,V) = [V, \nabla_V V] + dS_V\, V - dV\, S_V$.

PROOF. First we compute

$$\big(d^2V(V,Y)\big)_i = \sum_{j,k} \frac{\partial^2 v^i}{\partial x^j \partial x^k} v^j y^k = \sum_k y^k \frac{\partial}{\partial x^k} \sum_j \frac{\partial v^i}{\partial x^j} v^j - \sum_{j,k} \frac{\partial v^i}{\partial x^j} \frac{\partial v^j}{\partial x^k} y^k.$$

From this we deduce using properties (2.1) and (2.3)

$$\begin{aligned} d^2V(V,Y) =& d\big(dV\,V\big)Y - dV\,dV\,Y \\ =& d\big(\nabla_V V + S_V\big)Y - dV\big(\nabla_Y V + S(V,Y)\big) \\ =& \nabla_Y(\nabla_V V) + S(Y,\nabla_V V) + dS_V Y - \nabla_{\nabla_Y V} Y \\ & - S(V,\nabla_Y V) - dV\,S(V,Y). \end{aligned}$$

The result now follows by applying the formula (2.1). □

LEMMA 8.6.

$$\begin{aligned} d^3V(V,V,V) =& 2\nabla_V(\nabla_{\nabla_V V}V) - 2\nabla_{\nabla_V V}(\nabla_V V) + \nabla_{[V,\nabla_V V]}V - 2[V,\nabla_V(\nabla_V V)\,] \\ & + 3[V,[V,\nabla_V V]\,] + 2S(V,\nabla_V(\nabla_V V)) - 2S(\nabla_V V,\nabla_V V) \\ & - S(V,[V,\nabla_V V]) + 2dV\,S_\nabla + [V,[V,S_V]] - 2dS_V\nabla_V V \\ & + dV\,[V,S_V] - 2d^2V(V,S_V). \end{aligned}$$

PROOF. Using the identity

$$d^3V(V,V,V) = [V,[V,dV\,V]\,] + dV[V,dV\,V] - 2\,d^2V(V,dV\,V),$$

Lemma 8.5 and (2.3), we obtain the result. □

LEMMA 8.7. *Let $c : \mathbb{R} \mapsto M$ and $c(0) = p$; then*

$$\begin{aligned} c^{(4)}(0) =& d^3V(V,V,V) + 3\,d^2V(dV\,V,V) + dV\,d^2V(V,V) + dV\,dV\,dV\,V \\ =& \nabla_V(\nabla_V(\nabla_V V)) - \nabla_V(\nabla_{\nabla_V V}V) + \nabla_{\nabla_V V}\nabla_V V \\ & + \nabla_{[V,\nabla_V V]}V + 2S(V,[V,\nabla_V V]) + S(\nabla_V V,\nabla_V V) \\ & + d^2V(V,S_V) + [V,[V,S_V]\,] + 2dV\,dS_V\,V - dV\,dV\,S_V + dS_V\,\nabla_V V. \end{aligned}$$

PROOF. Using (2.1) and (2.3) as usual we obtain

$$\begin{aligned} dV\,dV\,\nabla_V V =& \nabla_V(\nabla_V(\nabla_V V)) - [V,\nabla_V(\nabla_V V)] - dV\,[V,\nabla_V V] \\ & + dV\,S_\nabla + S(V,\nabla_V(\nabla_V V)). \end{aligned}$$

Combining this with lemmas 8.5 and 8.6 gives the result. □

Acknowledgement.

The second author wants to thank Professor E.-M. Salonen at the Mechanics Laboratory of Helsinki University of Technology for helpful discussions concerning the final numerical example.

REFERENCES

1. U. Ascher and P. Lin, *Sequential regularization methods for nonlinear higher index DAEs*, SIAM J. Sci. Comp., 18 (1997), pp. 160–181.
2. M. Berz, C. Bischof, G. Corliss, and A. Griewank, eds., *Computational Differentiation*, SIAM, Philadelphia, PA, 1996.
3. J. Butcher, *The Numerical Solution of Ordinary Differential Equations*, Wiley, Chichester, 1987.
4. B. Char, K. Geddes, G. Bonnet, B. Leong, M. Monagan, and S. Watt, *Maple V Language Reference Manual*, Springer-Verlag, Berlin, 1991.
5. D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties and Algorithms*, Springer-Verlag, Berlin, 1992.
6. G.-M. Greuel, G. Pfister, and H. Schönemann, *Singular version 1.2 User Manual*, Reports on Computer Algebra, no. 21, Centre for Computer Algebra, University of Kaiserslautern, June 1998.
7. E. Hairer, S. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I, Nonstiff Problems*, Springer Series in Comp. Math., Vol. 8, Springer-Verlag, Berlin, 1987.
8. L. Petzold, Y. Ren, and T. Maly, *Regularization of higher-index DAEs with rank-deficient constraints*, SIAM J. Sci. Comp., 18 (1997), pp. 753–774.
9. D. Saunders, *The Geometry of Jet Bundles*, London Math. Soc., Lecture Notes Series, Vol. 142, Cambridge University Press, Cambridge, 1989.
10. J. M. Sanz Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, Chapman & Hall, London, 1994.
11. M. Spivak, *A Comprehensive Introduction to Differential Geometry, Vol. 1–5*, 2nd ed., Publish or Perish, 1979.
12. J. Tuomela, *On the numerical solution of involutive ordinary differential systems: Higher order methods*, Research Report A397, Helsinki University of Technology, Finland, 1998.
13. J. Tuomela and T. Arponen, *On the numerical solution of involutive ordinary differential systems: Introduction and numerical results*, IMA J. Numer. Anal., 20 (2000), pp. 561–599.
14. J. Tuomela and T. Arponen, *On the numerical solution of involutive ordinary differential systems: Higher order methods 2*, Research Report A409, Helsinki University of Technology, Finland, 1999.
15. S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer*, 2nd ed., Addison-Wesley, Redwood City, CA, 1991.