1. Introduction

We continue the work started in [4] 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 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 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 parctical 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 [4] 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 [4] 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. Actual numerical results using these methods will be presented separately.

2. Basic tools

We recall briefly the main notions that are needed. For more details on standard differential geometry we refer to [3] and on jets to [2]. 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 j'th 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 $\mathcal{E}$ 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.
- $\nabla_X (fY) = f\,\nabla_X Y + X(f)\,Y$

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 [4] 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

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

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 $\mathcal{D}$.*

Suppose that we are given a system of k q'th order ordinary differential equations

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

We interprete f as a morphism of bundles $J_q(\mathcal{E})$ and $\mathbb{R} \times \mathbb{R}^k$ which in terms of coordinates can be taken to be a map $\mathbb{R}^{(n+1)q+1} \to \mathbb{R}^k$. The above equation then defines a certain submanifold of $J_q(\mathcal{E})$ which we denote by $\mathcal{R}_q$ and in terms of coordinates $\mathcal{R}_q$ is simply given by $f^{-1}(0)$. 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.

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 let V be the vector field associated to it, i.e. $|V_p| = 1$ and $V_p \in \mathcal{D}_p$. Given a 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 $1 \leq k < n$, 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 $df_p\big(\tilde{c}^{(n)}(0) - c^{(n)}(0)\big) = 0$. $\square$

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 3.1. *For any n there are vectors $w^k \in TM_p$, $1 \leq k \leq n$, 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 $w^k \in TM_p$, $1 \leq k < n$, 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 $p^* = p + p^1 h + p^2 h^2 + \ldots$; 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 $q = p + q^1 h + q^2 h^2 + \ldots$; evidently $p^k = q^k$ for $1 \leq k < n$, 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

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

$\square$

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

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

Hence using lemma 5.1 we compute that $q^2 = \frac{1}{2}\, S(V_p, V_p)$ which combined with $c''(0) = dV\, V_p$ and (2.3) implies that

$$\tfrac{1}{2}\, c''(0) - q^2 = \tfrac{1}{2}\, (\nabla_{V_p} V)_p$$

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

Corollary 3.1. *If q is the solution of* (3.2) *with $w^1 = V_p$ and $w^2 = \frac{1}{2}(\nabla_{V_p}V)_p$, 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. To proceed in the computation of higher order terms 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.

In the computations that follow the formula (2.3) and lemma 5.1 are used very often. Note also that $\pi_n = (df)^t B^{-1} df$.

Proposition 3.1. *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

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

By theorem 3.1, 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 5.2 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 5.3 and 5.4 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 3.2. *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 3.1 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 3.2. *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

Lemma 3.3. *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. Expanding (3.1) we get $q^1 = V$ and $\mu^1 = B^{-1} df\, Y$. Then

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

Hence $B\mu^2 = df\big((dV)^t Y - \frac{1}{2}\, S_V\big)$ and

$$q^2 = \tfrac{1}{2}\, S_V + (dV)^t Y - (df)^t B^{-1} df\, (dV)^t Y = \tfrac{1}{2}\, S_V + \pi_t\big((dV)^t Y\big)$$

Consequently $c(h) - q = O(h^3)$, if $\pi_t\big((dV)^t Y\big) = \frac{1}{2}\, \nabla_V V$. Let $z^k \in TM_p$; then

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

For all $z^k \in TM_p$ we then get equations

$$\langle S(V, z^k), Y\rangle = \tfrac{1}{2}\, \langle \nabla_V V, z^k\rangle$$

From this the result follows. $\square$

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

Proposition 3.3. *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. $\square$

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 $c(h) - \hat{q} = O(h^3)$.

Let us recall the classical order conditions for Runge-Kutta methods, see for example [1]. Let $A \in \mathbb{R}^{r \times r}$ be the matrix containing the coefficients of the (r–stage) method. Let $b = (b_1, \ldots, b_r)$ and $\mathbf{1} = (1, \ldots, 1)$. Let us denote the component-wise (or Schur or Hadamard) product of vectors by $a \diamond b = (a_1 b_1, \ldots, a_r b_r)$. 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. $\square$

Next let us consider a scheme with 3 stages.

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

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

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

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

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

Comparing to lemma 5.3 we conclude that $q^3 = \frac{1}{6}\, c'''(0)$ and $\mu^3 = 0$. ☐

Let us then consider a scheme with 4 stages.

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

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 5.2 and 5.5

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

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

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

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

$$\begin{aligned} B\bar{\mu}^3 = & df \Big((u_5 u_6 - \tfrac{1}{2}\, u_5^3)(dV)^t S_V + (u_8 - \tfrac{1}{6}\, u_5^3) dS_V V + \\ & (a_{21}a_{32}a_{43} + \tfrac{1}{3}\, u_5^3 - u_5 u_6) S_\nabla \Big) \\ \bar{q}^3 = & u_7 [V, \nabla_V V] + a_{21}a_{32}a_{43} \nabla_V (\nabla_V V) + (u_5 u_6 - \tfrac{1}{3}\, u_5^3) S_\nabla + \\ & (u_8 + \tfrac{1}{2}\, u_5^3 - u_5 u_6) \pi_t (dS_V V) + \tfrac{1}{6}\, u_5^3 \pi_n (dS_V V) \end{aligned}$$

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

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

$$\begin{aligned} V_{\bar{q}} =& V_p + u_5 dV\, V\, h + \Big(u_6 \big(\nabla_V (\nabla_V V) + S_\nabla \big) + \\ & \big(\tfrac{1}{2}\, u_5^2 - u_6 \big) [V, \nabla_V V] + \tfrac{1}{2}\, u_5^2 dS_V V \Big) h^2 + \\ & \Big(dV\, \bar{q}^3 + u_5 u_6 d^2 V(V, \nabla_V V) + \\ & \tfrac{1}{2}\, u_5^3 d^2 V(V, S_V) + \tfrac{1}{6}\, u_5^3 d^3 V(V, V, V) \Big) h^3 + O(h^4) \end{aligned} \tag{4.11}$$

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

$$
\begin{aligned}
q^3 + (df_p)^t \mu^3 =& \big(\tfrac{1}{2} \langle b, A\mathbf{1} \diamond A\mathbf{1}\rangle - \langle b, A^2\mathbf{1}\rangle\big) [V, \nabla_V V] + \\
& \tfrac{1}{2} \langle b, A\mathbf{1} \diamond A\mathbf{1}\rangle dS_V V + \langle b, A^2\mathbf{1}\rangle \big(\nabla_V(\nabla_V V) + S_\nabla\big) \\
=& \tfrac{1}{6} \Big(\nabla_V(\nabla_V V) + dS_V\, V + S_\nabla\Big)
\end{aligned}
\tag{4.12}
$$

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

$$
q^4 + (df)^t \mu^4 = b_2 V^3_{\tilde{q}} + b_3 V^3_{\hat{q}} + b_4 V^3_{\bar{q}}
\tag{4.13}
$$

Anyway, at this point the computations became so tedious that it was necessary to use Mathematica [5]. Now guided by the previous steps we expect that $\mu^4 = 0$. Indeed, using the conditions (4.5) and simplifying we find that $b_2 V^3_{\tilde{q}} + b_3 V^3_{\hat{q}} + b_4 V^3_{\bar{q}} = \frac{1}{24} c^{(4)}(0)$, where $c^{(4)}(0)$ is given in lemma 5.8. There is one feature in this computation which is perhaps worth commenting. Looking at the right hand side of (4.13) there seems to be a term $dV\, S_\nabla$, but curiously its coefficient is zero *without* assuming the conditions (4.5). □

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

$$
\begin{cases}
\tilde{q} + (df_{\tilde{q}})^t \tilde{\mu} = p + h a_{11} V_{\tilde{q}} \\
f(\tilde{q}) = 0 \\
q + (df_q)^t \mu = p + h b_1 V_{\tilde{q}} \\
f(q) = 0
\end{cases}
\tag{4.14}
$$

Proposition 4.4. *The scheme* (4.14) *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

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

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.

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

Proposition 4.5. *The scheme* (4.15) *is of order 3 if and only if conditions* (4.2), (4.3) *and* (4.4) *are satisfied.*

Proof. We readily get $\tilde\mu^1 = \hat\mu^1 = 0$, $\tilde q^1 = (a_{11}+a_{12})V$ and $\hat q^1 = (a_{21}+a_{22})V$. The second order terms for $\tilde q$ give

$$
\begin{aligned} &\tilde q^2 + (df_p)^t \tilde\mu^2 = t_1 dV\,V \\ &df\,\tilde q^2 + \tfrac12 (a_{11}+a_{12})^2 d^2 f(V,V) = 0 \end{aligned}
$$

where $t_1 = a_{11}(a_{11}+a_{12}) + a_{12}(a_{21}+a_{22})$. Hence

$$
\begin{aligned} B\tilde\mu^2 =& df\Big(t_1 - \tfrac12 (a_{11}+a_{12})^2\Big) S_V \\ \tilde q^2 =& t_1 \nabla_V V + \tfrac12 (a_{11}+a_{12})^2 S_V \end{aligned}
$$

Similar computations show that

$$
\begin{aligned} B\hat\mu^2 =& df\Big(t_2 - \tfrac12 (a_{21}+a_{22})^2\Big) S_V \\ \hat q^2 =& t_2 \nabla_V V + \tfrac12 (a_{21}+a_{22})^2 S_V \end{aligned}
$$

where $t_2 = a_{21}(a_{11}+a_{12}) + a_{22}(a_{21}+a_{22})$. This implies

$$
(4.16)\qquad \begin{aligned} V_{\tilde q} =& V_p + (a_{11}+a_{12}) dV\,V\,h + \Big(t_1\big(\nabla_V(\nabla_V V) + S_\nabla\big) + \\ & \big(\tfrac12 (a_{11}+a_{12})^2 - t_1\big)[V, \nabla_V V] + \tfrac12 (a_{11}+a_{12})^2 dS_V V\Big) h^2 + O(h^3) \\ V_{\hat q} =& V_p + (a_{21}+a_{22}) dV\,V\,h + \Big(t_2\big(\nabla_V(\nabla_V V) + S_\nabla\big) + \\ & \big(\tfrac12 (a_{21}+a_{22})^2 - t_2\big)[V, \nabla_V V] + \tfrac12 (a_{21}+a_{22})^2 dS_V V\Big) h^2 + O(h^3) \end{aligned}
$$

Then expanding q we find that first that $\mu^1 = 0$ and $q^1 = (b_1+b_2)V$. Then the next equations are

$$
\begin{aligned} &q^2 + (df_p)^t \mu^2 = \langle b, A\mathbf{1}\rangle dV\,V \\ &df\,q^2 + \tfrac12 d^2 f(V,V) = 0 \end{aligned}
$$

which imply $\langle b, A\mathbf{1}\rangle = 1/2$ and $\mu^2 = 0$ as usual. The equation for q^3 is the same as (4.12). Hence $q^3 = \frac16 c'''(0)$ and $\mu^3 = 0$. □

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 [1].

Let us then consider the general three stage method.

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

Proposition 4.6. *The scheme* (4.17) *is of order 4 if and only if conditions* (4.2), (4.3), (4.4) *and* (4.5) *are satisfied.*

Proof. We readily get $\tilde{\mu}^1 = \hat{\mu}^1 = \bar{\mu}^1 = 0$, $\tilde{q}^1 = (a_{11} + a_{12} + a_{13})V = \tilde{r}_1 V$, $\hat{q}^1 = (a_{21} + a_{22} + a_{23})V = \hat{r}_1 V$ and $\bar{q}^1 = (a_{31} + a_{32} + a_{33})V = \bar{r}_1 V$. The second order terms for $\tilde{q}$ give

$$
\begin{aligned} \tilde{q}^2 + (df_p)^t \tilde{\mu}^2 &= \tilde{r}_2 dV\, V \\ df\, \tilde{q}^2 + \tfrac{1}{2}\, \tilde{r}_1^2 d^2 f(V, V) &= 0 \end{aligned}
$$

where $\tilde{r}_2 = a_{11}\tilde{r}_1 + a_{12}\hat{r}_1 + a_{13}\bar{r}_1$. Hence

$$
\begin{aligned} B\tilde{\mu}^2 =& df \left(\tilde{r}_2 - \tfrac{1}{2}\, \tilde{r}_1^2\right) S_V \\ \tilde{q}^2 =& \tilde{r}_2 \nabla_V V + \tfrac{1}{2}\, \tilde{r}_1^2 S_V \end{aligned}
$$

Similar computations show that

$$
\begin{aligned} B\hat{\mu}^2 =& df \left(\hat{r}_2 - \tfrac{1}{2}\, \hat{r}_1^2\right) S_V \\ \hat{q}^2 =& \hat{r}_2 \nabla_V V + \tfrac{1}{2}\, \hat{r}_1^2 S_V \\ B\bar{\mu}^2 =& df \left(\bar{r}_2 - \tfrac{1}{2}\, \bar{r}_1^2\right) S_V \\ \bar{q}^2 =& \bar{r}_2 \nabla_V V + \tfrac{1}{2}\, \bar{r}_1^2 S_V \end{aligned}
$$

where $\hat{r}_2 = a_{21}\tilde{r}_1 + a_{22}\hat{r}_1 + a_{23}\bar{r}_1$ and $\bar{r}_2 = a_{31}\tilde{r}_1 + a_{32}\hat{r}_1 + a_{33}\bar{r}_1$. To compute the third order terms we need the second order terms in the expansions of $V_{\tilde{q}}$, $V_{\hat{q}}$ and $V_{\bar{q}}$. They are as follows:

$$
\begin{aligned} V_{\tilde{q}}^2 =& \tilde{r}_2 \big(\nabla_V (\nabla_V V) + S_\nabla\big) + \big(\tfrac{1}{2}\, \tilde{r}_1^2 - \tilde{r}_2\big)[V, \nabla_V V] + \tfrac{1}{2}\, \tilde{r}_1^2 dS_V V \\ V_{\hat{q}}^2 =& \hat{r}_2 \big(\nabla_V (\nabla_V V) + S_\nabla\big) + \big(\tfrac{1}{2}\, \hat{r}_1^2 - \hat{r}_2\big)[V, \nabla_V V] + \tfrac{1}{2}\, \hat{r}_1^2 dS_V V \\ V_{\bar{q}}^2 =& \bar{r}_2 \big(\nabla_V (\nabla_V V) + S_\nabla\big) + \big(\tfrac{1}{2}\, \bar{r}_1^2 - \bar{r}_2\big)[V, \nabla_V V] + \tfrac{1}{2}\, \bar{r}_1^2 dS_V V \end{aligned}
$$

Then the third order terms give

$$
\begin{aligned} &\tilde{q}^3 + (df_p)^t \tilde{\mu}^3 + \tilde{r}_1 d^2 f(V, \cdot)\tilde{\mu}^2 = \tilde{r}_3 \big(\nabla_V (\nabla_V V) + S_\nabla\big) + \tilde{r}_4 [V, \nabla_V V] + \tilde{r}_5 dS_V V \\ &df\, \tilde{q}^3 + \tilde{r}_2 d^2 f(V, \nabla_V V) + \tfrac{1}{2}\, \tilde{r}_1^3 d^2 f(V, S_V) + \tfrac{1}{6}\, \tilde{r}_1^3 d^3 f(V, V, V) = 0 \end{aligned}
$$

where

$$\begin{aligned}
\tilde r_3 =& a_{11}\tilde r_2 + a_{12}\hat r_2 + a_{13}\bar r_2\\
\tilde r_4 =& a_{11}\big(\tfrac12\,\tilde r_1^2 - \tilde r_2\big) + a_{12}\big(\tfrac12\,\hat r_1^2 - \hat r_2\big) + a_{13}\big(\tfrac12\,\bar r_1^2 - \bar r_2\big)\\
\tilde r_5 =& \tfrac12\,a_{11}\tilde r_1^2 + \tfrac12\,a_{12}\hat r_1^2 + \tfrac12\,a_{13}\bar r_1^2
\end{aligned}$$

Solving these yields

$$\begin{aligned}
B\tilde\mu^3 &= df\Big((\tilde r_1\tilde r_2 - \tfrac12\,\tilde r_1^3)(dV)^t S_V + (\tilde r_5 - \tfrac16\,\tilde r_1^3)dS_V V +\\
&\qquad (\tilde r_3 + \tfrac13\,\tilde r_1^3 - \tilde r_1\tilde r_2)S_\nabla\Big)\\
\tilde q^3 &= \tilde r_4[V,\nabla_V V] + \tilde r_3\nabla_V(\nabla_V V) + (\tilde r_1\tilde r_2 - \tfrac13\,\tilde r_1^3)S_\nabla +\\
&\qquad (\tilde r_5 + \tfrac12\,\tilde r_1^3 - \tilde r_1\tilde r_2)\pi_t(dS_V V) + \tfrac16\,\tilde r_1^3\pi_n(dS_V V)
\end{aligned}$$

In exactly the same way we obtain a formula for $\hat q^3$ (resp. $\bar q^3$) by replacing $\tilde r_i$ by $\hat r_i$ (resp. $\bar r_i$) using the formulas

$$\begin{aligned}
\hat r_3 =& a_{21}\tilde r_2 + a_{22}\hat r_2 + a_{23}\bar r_2\\
\hat r_4 =& a_{21}\big(\tfrac12\,\tilde r_1^2 - \tilde r_2\big) + a_{22}\big(\tfrac12\,\hat r_1^2 - \hat r_2\big) + a_{23}\big(\tfrac12\,\bar r_1^2 - \bar r_2\big)\\
\hat r_5 =& \tfrac12\,a_{21}\tilde r_1^2 + \tfrac12\,a_{22}\hat r_1^2 + \tfrac12\,a_{23}\bar r_1^2\\
\bar r_3 =& a_{31}\tilde r_2 + a_{32}\hat r_2 + a_{33}\bar r_2\\
\bar r_4 =& a_{31}\big(\tfrac12\,\tilde r_1^2 - \tilde r_2\big) + a_{32}\big(\tfrac12\,\hat r_1^2 - \hat r_2\big) + a_{33}\big(\tfrac12\,\bar r_1^2 - \bar r_2\big)\\
\bar r_5 =& \tfrac12\,a_{31}\tilde r_1^2 + \tfrac12\,a_{32}\hat r_1^2 + \tfrac12\,a_{33}\bar r_1^2
\end{aligned}$$

Hence we get

$$\begin{aligned}
V_{\tilde q}^3 =& dV\,\tilde q^3 + \tilde r_1\tilde r_2 d^2V(V,\nabla_V V) + \tfrac12\,\tilde r_1^3 d^2V(V,S_V) + \tfrac16\,\tilde r_1^3 d^3V(V,V,V)\\
V_{\hat q}^3 =& dV\,\hat q^3 + \hat r_1\hat r_2 d^2V(V,\nabla_V V) + \tfrac12\,\hat r_1^3 d^2V(V,S_V) + \tfrac16\,\hat r_1^3 d^3V(V,V,V)\\
V_{\bar q}^3 =& dV\,\bar q^3 + \bar r_1\bar r_2 d^2V(V,\nabla_V V) + \tfrac12\,\bar r_1^3 d^2V(V,S_V) + \tfrac16\,\bar r_1^3 d^3V(V,V,V)
\end{aligned}$$

Then expanding q we find first that $\mu^1 = 0$ and $q^1 = (b_1+b_2+b_3)V$. The next equations are

$$\begin{aligned}
q^2 + (df_p)^t\mu^2 &= \langle b, A\mathbf{1}\rangle dV\,V\\
df\,q^2 + \tfrac12\,d^2f(V,V) &= 0
\end{aligned}$$

which imply $\langle b, A\mathbf{1}\rangle = 1/2$ and $\mu^2 = 0$ as usual. The equation for q^3 is again the same as (4.12). Hence $q^3 = \frac16\,c'''(0)$ and $\mu^3 = 0$. Finally for the fourth order terms it was again necessary to use Mathematica. The equation is analogous to (4.13), namely

$$q^4 + (df)^t\mu^4 = b_1V_{\tilde q}^3 + b_2V_{\hat q}^3 + b_3V_{\bar q}^3 \tag{4.18}$$

and again we find that if the conditions (4.5) hold, the right hand side of (4.18) is $\frac{1}{24}\,c^{(4)}(0)$. Hence $q^4 = \frac{1}{24}\,c^{(4)}(0)$ and $\mu^4 = 0$. Again the coefficient of the term $dV\,S_\nabla$ on the right hand side of (4.18) vanishes without assuming the conditions (4.5). □

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: $a_{21} = 5/24$, $a_{22} = 1/3$, $a_{23} = -1/24$, $a_{31} = 1/6$, $a_{32} = 2/3$, $a_{33} = 1/6$.

5. 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 5.1. *Let $p \in M$ and $Y \in \Gamma(TM)$; then*

$$S(V_p, Y_p) = -(df_p)^t B^{-1} d^2 f(V_p, Y_p)$$

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. □

Lemma 5.2. *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

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

since $\langle df^j, V\rangle = 0$. The proof of the other statement is similar. □

Lemma 5.3. *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 5.4.

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

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 5.5.

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

Proof.

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

Hence

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

which combined with lemma 5.2 gives the result. □

Lemma 5.6. *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 5.7.

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

Proof. Using the identity $d^3V(V,V,V) = [V,[V,dV\, V]] + dV[V, dV\, V] - 2\, d^2V(V, dV\, V)$, lemma 5.6 and (2.3) we obtain the result. □

Lemma 5.8. *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 5.6 and 5.7 gives the result. □

References

1. J. Butcher, *The numerical solution of ordinary differential equations*, Wiley, 1987.
2. D. Saunders, *The geometry of jet bundles*, London Math. Soc. Lecture note series, vol. 142, Cambridge university press, 1989.
3. M. Spivak, *A comprehensive introduction to differential geometry, vol 1 - 5, 2nd ed.*, Publish or Perish, 1979.
4. J. Tuomela and T. Arponen, *On the numerical solution of involutive ordinary differential systems: Introduction and numerical results*, Research Report A393, Helsinki University of Technology, 1998.
5. S. Wolfram, *Mathematica: a system for doing mathematics by computer, 2nd ed.*, Addison-Wesley, 1991.