# Order theory for discrete gradient methods

Sølve Eidnes*

March 23, 2020

**Abstract:** We present a subclass of the discrete gradient methods, which are integrators designed to preserve invariants of ordinary differential equations. From a formal series expansion of the methods, we derive conditions for arbitrarily high order. We devote considerable space to the average vector field discrete gradient, from which we get P-series methods in the general case, and B-series methods for canonical Hamiltonian systems. Higher order schemes are presented and applied to the Hénon–Heiles system and a Lotka–Volterra system.

**Keywords:** Geometric integration, discrete gradients, energy preservation, B-series, order of accuracy, high-order methods

**AMS subject classification (2010):** 65L05, 65P10, 37M15, 05C05

## 1 Energy preservation and discrete gradient methods

For an ordinary differential equation (ODE)

$$\dot{x} = f(x), \quad x \in \mathbb{R}^d, \quad f : \mathbb{R}^d \to \mathbb{R}^d, \tag{1.1}$$

a first integral, or invariant, is a function $H : \mathbb{R}^d \to \mathbb{R}$ such that $H(x(t)) = H(x(t_0))$ along the solution curves of (1.1). If we can write

$$f(x) = S(x)\nabla H(x), \tag{1.2}$$

where $S(x) : \mathbb{R}^{d\times d} \to \mathbb{R}^d$ is a skew-symmetric matrix, then (1.1) preserves $H$: this follows from the skew-symmetry of $S(x)$, which yields

$$\frac{\mathrm{d}}{\mathrm{d}t} H(x) = \nabla H(x)^T \dot{x} = \nabla H(x)^T S(x)\nabla H(x) = 0. \tag{1.3}$$

The converse is also true: McLachlan et al. showed in [20] that, whenever (1.1) has a first integral $H$, there exists a skew-symmetric matrix $S(x)$, bounded near every non-degenerate critical point of $H$, such that (1.1) can be written on what is called the *skew-gradient form*:

$$\dot{x} = S(x)\nabla H(x). \tag{1.4}$$

*Department of Mathematical Sciences, Norwegian University of Science and Technology, N–7491 Trondheim
E-mail: `solve.eidnes@ntnu.no`

The proof provided in [20] for this is based on presenting a general form of one such $S(x)$, the so-called default formula

$$S(x) = \frac{f(x)\nabla H(x)^T - \nabla H(x) f(x)^T}{\nabla H(x)^T \nabla H(x)}. \tag{1.5}$$

Unless $d = 2$, this is generally not a unique choice of $S(x)$, as e.g.

$$S(x) = \frac{f(x) g(x)^T - g(x) f(x)^T}{g(x)^T \nabla H(x)}$$

will satisfy (1.2) for any non-vanishing function $g : \mathbb{R}^d \to \mathbb{R}^d$. Many ODEs with first integrals have a well-known skew-gradient form (1.4). This includes Poisson systems, and the important class consisting of canonical Hamiltonian ODEs. For the latter, $S$ will be constant, so that we may write

$$\dot{x} = S \nabla H(x). \tag{1.6}$$

A numerical integrator preserving a first integral $H$ exactly is called an integral-preserving, or *energy-preserving*, method. Starting in the late 1970s, a few energy-preserving methods were proposed which relied on some discrete analogue of the property (1.3), see e.g. [4, 15, 17, 16]. Most prominent among these is the class of methods called discrete gradient methods, defined formally by Gonzalez in [11] and given their current name in [20].

Given the first integral $H$, a discrete gradient $\overline{\nabla} H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is a function satisfying the conditions

$$\overline{\nabla} H(x, y)^{\mathrm{T}} (y - x) = H(y) - H(x), \tag{1.7}$$

$$\overline{\nabla} H(x, x) = \nabla H(x), \tag{1.8}$$

for all $x, y \in \mathbb{R}^d$. Introducing also the discrete approximation $\overline{S}(x, y, h)$ to $S(x)$, skew-symmetric and satisfying $\overline{S}(x, x, 0) = S(x)$, the corresponding discrete gradient method is given by

$$\frac{\hat{x} - x}{h} = \overline{S}(x, \hat{x}, h) \overline{\nabla} H(x, \hat{x}). \tag{1.9}$$

This scheme satisfies a discrete analogue to (1.3):

$$H(\hat{x}) - H(x) = h \overline{\nabla} H(x, \hat{x})^T \overline{S}(x, \hat{x}, h) \overline{\nabla} H(x, \hat{x}) = 0.$$

We say that (1.9) is consistent to the skew-gradient system (1.4), since $\overline{S}(x, \hat{x}, h)$ is a consistent approximation of $S(x)$ and $\overline{\nabla} H(x, \hat{x})$ is a consistent approximation of $\nabla H(x)$.

If $d \geq 2$, there are in general infinitely many functions satisfying (1.7)–(1.8). Many explicit definitions of concrete discrete gradients have been suggested, and we will discuss the most prominent among them in Section 2.1. One of these is the *average vector field (AVF) discrete gradient*, first introduced in [14] and sometimes called the mean value discrete gradient [20]. For a given $H$, it is given by the average of $\nabla H$ on the segment $[x, y]$:

$$\overline{\nabla}_{\mathrm{AVF}} H(x, y) = \int_0^1 \nabla H((1 - \xi) x + \xi y) \, \mathrm{d}\xi. \tag{1.10}$$

When applied to the constant $S$ system (1.6), the discrete gradient method with $\overline{S}(x,y,h)=S$ and $\overline{\nabla} H=\overline{\nabla}_{\mathrm{AVF}} H$ coincides with the scheme

$$\frac{\hat{x}-x}{h}=\int_0^1 f((1-\xi)x+\xi\hat{x})\,\mathrm{d}\xi. \tag{1.11}$$

This is sometimes viewed as a method by itself, applicable to any system (1.1), in which case it is called the average vector field (AVF) method [26]. This was shown in [2] to be a B-series method.

As pointed out in [20], the discrete gradient is restricted by its definition to be at best a second order approximation to point values of $\nabla H$. In much of the literature on discrete gradient methods, see e.g. [11, 13], the approximation $\overline{S}$ is defined as being independent of $h$. In that case, the discrete gradient scheme (1.9) can at best guarantee second order convergence towards the exact solution. Over the last two decades, there have been published some notable papers on higher order discrete gradient methods. McLaren and Quispel were first out with their bootstrapping technique derived in [21, 22]. Given any discrete gradient $\overline{\nabla} H$ and an approximation to $S(x)$ given by $\overline{S}(x,y,h)$, they compare the Taylor expansion of the corresponding discrete gradient scheme to that of the exact solution, and thus find a new approximation $\tilde{S}(x,y,h)$ to $S(x)$ which yields higher order. This quickly becomes a very involved procedure, but by using a symmetric discrete gradient, they derive fourth order methods. A downside of this method is that the schemes of order higher than two require the calculation of tensors of order three or higher at every time step.

A fourth order generalization of the AVF method is proposed by the same authors in [26]. This can be viewed as a fourth-order discrete gradient method for all skew-gradient systems where $S$ is constant. Also worth mentioning in this setting is the collocation-like method introduced by Hairer [12] and then generalized to Poisson systems by Cohen and Hairer [5]. This is a multi-stage extension of the AVF discrete gradient method. To get higher than second order, more than one stage is required. In that case the method is not a discrete gradient method, although it is energy-preserving.

Norton et al. show in [24] that linear projection methods can be viewed as a class of discrete gradient methods for skew-gradient systems with $S(x)$ given by the default formula (1.5). In connection to this, Norton and Quispel suggest in [25] the class of approximations to (1.5) given by

$$\overline{S}(x,y,h)=\frac{\tilde{f}(x,y,h)\tilde{g}(x,y,h)^T-\tilde{g}(x,y,h)\tilde{f}(x,y,h)^T}{\hat{g}(x,y,h)^T\breve{g}(x,y,h)}, \tag{1.12}$$

where $\tilde{f}(x,y,h)$ is a consistent approximation to $f(x)$, and $\tilde{g}(x,y,h)$, $\hat{g}(x,y,h)$ and $\breve{g}(x,y,h)$ are all consistent approximations to $\nabla H(x)$. The corresponding discrete gradient method then inherits the order of the method $\hat{x}=x+h\tilde{f}(x,\hat{x},h)$.

To the best of our knowledge, no one has so far suggested higher than fourth order discrete gradient methods for a general skew-gradient system (1.4). Furthermore, for this general case, all discrete gradient methods suggested of higher than second order involve tensors of order three or higher. Our aim with this paper is to remedy this. Largely inspired by the above mentioned references, especially [26, 21, 22], we present here a general form giving a class of approximations $\overline{S}(x,y,h)$ to any $S(x)$ in (1.4), with corresponding conditions for achieving an arbitrary order of the discrete gradient method (1.9). We do this step by step. In the next chapter, we derive some useful properties of a general discrete gradient and discuss the most common specific discrete gradients. Then we consider the AVF method and use order theory for B-series methods to obtain a generalization of this, with corresponding order conditions. In Chapter 4, we build on this to develop higher order discrete gradient methods for a general skew-gradient system, using the AVF discrete gradient. Then, in Chapter 5, we generalize this further to allow for a free choice of the discrete gradient, thus arriving at the general form $\overline{S}(x, y, h)$ mentioned above, and a formal series expansion of the corresponding discrete gradient methods. We present several examples of higher order schemes for the different cases, and conclude the paper with some numerical experiments.

## 2 A preliminary analysis of discrete gradients

To simplify notation in the following derivations, we define $g := \nabla H$. Furthermore, we suppress the first argument of $\overline{\nabla} H$ and define $\bar{g}(y) := \overline{\nabla} H(x, y)$. We use Einstein summation convention and write $\bar{g}(y)^i_j := \frac{\partial \bar{g}(y)^i}{\partial y^j}$ and so forth. Taylor expanding $\bar{g}(y)$ around $x$, we get

$$\begin{aligned}\bar{g}(y)^i = \bar{g}(x)^i + \bar{g}(x)^i_j(y^j - x^j) + \frac{1}{2}\bar{g}(x)^i_{jk}(y^j - x^j)(y^k - x^k)\\ + \frac{1}{6}\bar{g}(x)^i_{jkl}(y^j - x^j)(y^k - x^k)(y^l - x^l) + \mathcal{O}(|y - x|^4),\end{aligned} \tag{2.1}$$

or

$$\bar{g}(y) = \sum_{\kappa=0}^{\infty} \frac{1}{\kappa!}\bar{g}^{(\kappa)}(x)(y - x)^{\kappa}. \tag{2.2}$$

By the consistency criterion (1.8), we have $\bar{g}(x) = g(x)$. However, if we require the discrete gradient to be a differentiable function in its second argument, (1.8) follows directly from (1.7). To see this, we write (1.7) as

$$H(y) - H(x) = \bar{g}(y)_i(y^i - x^i). \tag{2.3}$$

Differentiating this with respect to $y^j$, we get

$$g(y)_j = H(y)_j = \bar{g}(y)_{i,j}(y^i - x^i) + \bar{g}(y)_j, \tag{2.4}$$

where $H_j = \frac{\partial H}{\partial y^j}$ and $\bar{g}(y)_{i,j} = \frac{\partial \bar{g}(y)_i}{\partial y^j}$. The case $y = x$ immediately gives $g(x)_j = \bar{g}(x)_j$, or (1.8). Assuming further that $\overline{\nabla} H \in C^2(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$, we can differentiate once more to get

$$g(y)_{j,k} = H(y)_{jk} = \bar{g}(y)_{i,jk}(y^i - x^i) + \bar{g}(y)_{j,k} + \bar{g}(y)_{k,j}, \tag{2.5}$$

which means that

$$g(x)_{j,k} = H(x)_{jk} = \bar{g}(x)_{j,k} + \bar{g}(x)_{k,j},$$

or

$$\nabla^2 H(x) = D_2\overline{\nabla} H(x, x) + (D_2\overline{\nabla} H(x, x))^{\mathrm{T}}, \tag{2.6}$$

where $\nabla^2 H := D\nabla H$ denotes the Hessian of $H$, and $D_2\overline{\nabla} H$ denotes the Jacobian of $\overline{\nabla} H$ with respect to its second argument.

**Lemma 2.1.** *If the discrete gradient $\overline{\nabla}H$ is symmetric, i.e. $\overline{\nabla}H(x,y) = \overline{\nabla}H(y,x)$ for all $x, y \in \mathbb{R}^d$, then*

$$D_2\overline{\nabla}H(x,x) = \frac{1}{2}\nabla^2 H(x). \tag{2.7}$$

*Proof.* Disclosing the suppressed argument $x$ in (2.4), we have

$$g(y)_j = \frac{\partial}{\partial y^j}(\bar{g}(x,y)_i)(y^i - x^i) + \bar{g}(x,y)_j,$$

which we can differentiate by $x^k$ to get

$$0 = \frac{\partial^2}{\partial x^k \partial y^j}(\bar{g}(x,y)_i)(y^i - x^i) - \frac{\partial}{\partial y^j}\bar{g}(x,y)_k + \frac{\partial}{\partial x^k}\bar{g}(x,y)_j.$$

If $\overline{\nabla}H$ is symmetric,

$$\frac{\partial}{\partial x^k}\bar{g}(x,y)_j = \frac{\partial}{\partial x^k}\bar{g}(y,x)_j.$$

Thus, for $y = x$ we get $\bar{g}(x)_{k,j} = \bar{g}(x)_{j,k}$, or $(D_2\overline{\nabla}H(x,x))^{\mathrm{T}} = D_2\overline{\nabla}H(x,x)$. Inserting that in (2.6), we obtain (2.7). □

**Definition 2.1.** Given a discrete gradient $\overline{\nabla}H \in C^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$, we define the function $Q : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ by

$$Q(x,y) := \frac{1}{2}\left((D_2\overline{\nabla}H(x,y))^T - D_2\overline{\nabla}H(x,y)\right). \tag{2.8}$$

Note that $Q(x,y)$ is a skew-symmetric matrix. From (2.6), we see that $Q(x,x) = \frac{1}{2}\nabla^2 H(x) - D_2\overline{\nabla}H(x,x)$. Differentiating (2.8) with respect to the second argument and setting $y = x$, we obtain

$$(D_2Q(x,x))_{jkl} = \frac{1}{2}\bar{g}(x)_{k,jl} - \frac{1}{2}\bar{g}(x)_{j,kl}.$$

Similarly, differentiating (2.5) with respect to the second argument and setting $y = x$, we obtain

$$g(x)_{j,kl} = \bar{g}(x)_{j,kl} + \bar{g}(x)_{k,jl} + \bar{g}(x)_{l,jk}.$$

Using these results, we get that, for any $v \in \mathbb{R}^d$,

$$\begin{aligned}
(D_2Q(x,x)(v,v))_j = (D_2Q(x,x))_{jkl}v^k v^l &= \frac{1}{2}\bar{g}(x)_{k,jl}v^k v^l - \frac{1}{2}\bar{g}(x)_{j,kl}v^k v^l \\
&= \frac{1}{4}\bar{g}(x)_{k,jl}v^k v^l + \frac{1}{4}\bar{g}(x)_{l,jk}v^k v^l + \frac{1}{4}\bar{g}(x)_{j,kl}v^k v^l - \frac{3}{4}\bar{g}(x)_{j,kl}v^k v^l \\
&= \frac{1}{4}g(x)_{j,kl}v^k v^l - \frac{3}{4}\bar{g}(x)_{j,kl}v^k v^l,
\end{aligned}$$

or

$$D_2Q(x,x)(v,v) = \frac{1}{4}D^2\nabla H(x)(v,v) - \frac{3}{4}D_2^2\overline{\nabla}H(x,x)(v,v). \tag{2.9}$$

Continuing in this manner, we get the following general result, which will be useful when developing higher order discrete gradient methods.

**Lemma 2.2.** *For a discrete gradient $\overline{\nabla} H \in C^p(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R})$ and the corresponding $Q$ given by* (2.8),

$$D_2^\kappa \overline{\nabla} H(x,x) v^\kappa = \frac{1}{\kappa+1} D^\kappa \nabla H(x) v^\kappa - \frac{2\kappa}{\kappa+1} D_2^{\kappa-1} Q(x,x) v^\kappa \quad \textit{for any } \kappa \in [1,p], v \in \mathbb{R}^d.$$

*Proof.* Differentiating (2.5) $\kappa - 1$ times by $y$ and setting $y = x$, we find that the $\kappa$-th derivatives of $g(x)$ can be expressed by the $\kappa$-th derivatives of $\bar{g}(x)$ through the relation

$$g(x)_{j,I} = \bar{g}(x)_{j,I} + \sum_{m=1}^{\kappa} \bar{g}(x)_{i_m,\{j,I_m\}}, \qquad \text{for all } j, I, \kappa, \tag{2.10}$$

where $I = \{i_1, i_2, \ldots, i_\kappa\}$ is an ordered set of $\kappa$ indices, and $I_m = I \backslash \{i_m\} = \{i_1, i_2, \ldots i_{m-1}, i_{m+1}, \ldots, i_\kappa\}$, i.e. $I$ with the $m$-th element excluded. Similarly, by continued differentiation of (2.8), we obtain

$$(D_2^{\kappa-1} Q(x,x))_{j,I} = \frac{1}{2} \bar{g}(x)_{i_1,\{j,I_1\}} - \frac{1}{2} \bar{g}(x)_{j,I}.$$

Thus

$$\begin{aligned}
(D_2^{\kappa-1} Q(x,x) v^\kappa)_j &= (D_2^{\kappa-1} Q(x,x))_{j,I} v^I = \frac{1}{2} \bar{g}(x)_{i_1,\{j,I_1\}} v^{\{j,I_1\}} - \frac{1}{2} \bar{g}(x)_{j,I} v^I \\
&= \frac{1}{2\kappa} \sum_{m=1}^{\kappa} \bar{g}(x)_{i_m,\{j,I_m\}} v^{\{j,I_m\}} + \frac{1}{2\kappa} \bar{g}(x)_{j,I} v^I - \frac{1}{2\kappa} \bar{g}(x)_{j,I} v^I - \frac{1}{2} \bar{g}(x)_{j,I} v^I \\
&= \frac{1}{2\kappa} g(x)_{j,I} v^I - \left(\frac{1}{2\kappa} + \frac{1}{2}\right) \bar{g}(x)_{j,I} v^I = \frac{1}{2\kappa} g(x)_{j,I} v^I - \frac{\kappa+1}{2\kappa} \bar{g}(x)_{j,I} v^I.
\end{aligned}$$

□

## 2.1 A review of explicitly defined discrete gradients

While introducing the discrete gradient methods in [11], Gonzalez also gave an example of a discrete gradient satisfying (1.7)–(1.8): the *midpoint discrete gradient* is given by

$$\overline{\nabla}_{\mathrm{M}} H(x,y) := \nabla H\left(\frac{x+y}{2}\right) + \frac{H(y) - H(x) - \nabla H\left(\frac{x+y}{2}\right)^T (y-x)}{(y-x)^T (y-x)} (y-x).$$

Even when $H$ is analytical, this discrete gradient is often not; the second order partial derivatives are in general singular in $y = x$. For that reason, it is not suited for achieving higher order methods by the techniques we consider in this paper.

The *Itoh–Abe discrete gradient*, introduced in [15], notably does not require evaluation of the gradient. This discrete gradient, which has also been called the coordinate increment discrete gradient [20], is defined by

$$\overline{\nabla}_{\mathrm{IA}} H(x,y) := \sum_{j=1}^{d} \alpha_j e_j, \tag{2.11}$$

where $e_j$ is the $j^{\text{th}}$ canonical unit vector and

$$\alpha_j = \begin{cases} \dfrac{H(w_j) - H(w_{j-1})}{y^j - x^j} & \text{if } y^j \neq x^j, \\ \frac{\partial H}{\partial x^j}(w_{j-1}) & \text{if } y^j = x^j, \end{cases}$$
$$w_j = \textstyle\sum_{i=1}^{j} y^i e_i + \sum_{i=j+1}^{n} x^i e_i.$$

While the other discrete gradients we consider in this paper are symmetric and thus second order approximations to $\nabla H$, the Itoh–Abe discrete gradient is only of first order. However, a second order discrete gradient, which we call the *symmetrized Itoh–Abe (SIA) discrete gradient*, is given by

$$\overline{\nabla}_{\mathrm{SIA}} H(x,y) := \frac{1}{2}\left(\overline{\nabla}_{\mathrm{IA}} H(x,y) + \overline{\nabla}_{\mathrm{IA}} H(y,x)\right). \tag{2.12}$$

Furihata presented the discrete variational derivative method for a class of partial differential equations (PDEs) in [9], a method which has been developed further by Furihata, Matsuo and co-authors in a series of papers, e.g. [19, 27], as well as the monograph [10]. As shown in [7], these schemes can also be obtained by semi-discretizing the PDE in space and then applying a discrete gradient method on the resulting system of ODEs. The specific discrete gradient that gives the schemes of Furihata and co-authors is defined for a class of invariants that includes all polynomial functions:

**Definition 2.2.** Assume that we can write the first integral as

$$H(x) = \sum_l c_l \prod_{j=1}^{d} f_j^l(x^j), \tag{2.13}$$

for functions $f_j^l : \mathbb{R} \to \mathbb{R}$. The *Furihata discrete gradient* $\overline{\nabla}_{\mathrm{F}} H(x,y)$ is defined by

$$\overline{\nabla}_{\mathrm{F}} H(x,y) := \sum_{j=1}^{d} \alpha_j e_j, \tag{2.14}$$

where $e_j$ is the $j^{\text{th}}$ canonical unit vector and

$$\alpha_j = \begin{cases} \sum\limits_l \frac{c_l}{2} \frac{f_j^l(y^j) - f_j^l(x^j)}{y^j - x^j} \left( \prod\limits_{k=1}^{j-1} f_k^l(x^k) + \prod\limits_{k=1}^{j-1} f_k^l(y^k) \right) \prod\limits_{k=j+1}^{d} \frac{f_k^l(x^k) + f_k^l(y^k)}{2} & \text{if } y^j \neq x^j, \\ \sum\limits_l \frac{c_l}{2} \frac{\mathrm{d} f_j^l(x^j)}{\mathrm{d} x^j} \left( \prod\limits_{k=1}^{j-1} f_k^l(x^k) + \prod\limits_{k=1}^{j-1} f_k^l(y^k) \right) \prod\limits_{k=j+1}^{d} \frac{f_k^l(x^k) + f_k^l(y^k)}{2} & \text{if } y^j = x^j. \end{cases}$$

Lastly we consider the AVF discrete gradient (1.10), which distinguishes itself from the others in a number of ways.

**Lemma 2.3.** *The $Q(x,y)$ corresponding to the AVF discrete gradient is the zero matrix, since $(D_2 \overline{\nabla}_{AVF} H(x,y))^T = D_2 \overline{\nabla}_{AVF} H(x,y)$.*

*Proof.* For $\bar{g}(y) := \overline{\nabla}_{\mathrm{AVF}} H(x,y)$, we have

$$\begin{aligned}\bar{g}(y)_{i,j} &= \frac{\partial}{\partial y^j}\int_0^1 g((1-\xi)x+\xi y)_i \,\mathrm{d}\xi = \int_0^1 \frac{\partial}{\partial y^j} g((1-\xi)x+\xi y)_i \,\mathrm{d}\xi \\ &= \int_0^1 \xi g((1-\xi)x+\xi y)_{i,j}\,\mathrm{d}\xi = \int_0^1 \xi g((1-\xi)x+\xi y)_{j,i}\,\mathrm{d}\xi \\ &= \bar{g}(y)_{j,i}.\end{aligned}$$

□

**Proposition 2.4.** *The AVF discrete gradient is the unique discrete gradient satisfying* $(D_2\overline{\nabla}H(x,y))^T = D_2\overline{\nabla}H(x,y)$ *for all H, x and y, and it has the formal expansion*

$$\overline{\nabla}_{AVF} H(x,y) = \sum_{\kappa=0}^{\infty} \frac{1}{(\kappa+1)!} D^{\kappa}\nabla H(x)(y-x)^{\kappa}. \tag{2.15}$$

*Proof.* Assume that $\nabla H$ is an analytic function. As in the proof of Lemma 2.2, let $I = \{i_1, i_2, \ldots, i_\kappa\}$ be an ordered set of $\kappa$ indices, and let $I_m$ be $I$ with the $m^{\text{th}}$ element excluded. If $\bar{g}(y)_{i,j} = \bar{g}(y)_{j,i}$ for all $i, j$, then also

$$\bar{g}(y)_{i,I} = \bar{g}(y)_{i_m,\{i,I_m\}} \quad \text{for all } i, I, m. \tag{2.16}$$

Inserting (2.16) in (2.10) we get $g^{(\kappa)}(x) = (1+\kappa)\bar{g}^{(\kappa)}(x)$. Then inserting this for $\bar{g}^{(\kappa)}(x)$ in (2.2), we get (2.15), which uniquely defines the AVF discrete gradient. □

A consequence of the above result is that the AVF discrete gradient is the unique discrete gradient for which the scheme (1.9) with $\overline{S}(x,\hat{x},h) = S$ is a B-series method when applied to the system (1.6). Furthermore, from the Integrability Lemma (see e.g. [13, Lemma VI.2.7]) and the above, we have that it is the only discrete gradient which defines a gradient vector field in general:

**Corollary 2.4.1.** *The AVF discrete gradient is the gradient with respect to the second argument of a function* $\tilde{H}(x,y)$*. That is,*

$$\overline{\nabla}_{AVF} H(x,y) = \nabla_2 \tilde{H}(x,y),$$

*for some* $\tilde{H} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *and all* $x, y \in \mathbb{R}^d$*. The AVF discrete gradient is the unique discrete gradient to have this property for all H.*

As we see from the above definitions and discussion, each of the discrete gradients have their advantages and disadvantages. Gonzalez' midpoint discrete gradient is easily calculated from the energy $H$ and the gradient $\nabla H$, but it is in general only once differentiable. The Itoh–Abe discrete gradient does not require knowledge of the gradient, but is only a first order approximation of the gradient. The AVF discrete gradient is the unique discrete gradient whose series expansion is given by the differentials of the gradient. It does however require an integral to be calculated. If that poses a challenge, the SIA or Furihata discrete gradients are second-order alternatives, but the latter is only defined for $H$ of the form (2.13).

## 2.2 Third and fourth order schemes for the constant $S$ case

Consider now only the cases where $S$ is constant, i.e. (1.6). By comparing the Taylor series of the exact solution and that of the discrete gradient method, and by using the properties of the discrete gradient developed above, we may achieve higher order discrete gradient methods.

In search of a third order scheme, we assume that $\hat{x}$ is a third order in $h$ approximation of $x(t_0+h)$, and find

$$\begin{aligned}
S\overline{\nabla}H(x,\hat{x}) =& S(\nabla H(x) + D_2\overline{\nabla}H(x,x)(hS\nabla H(x) + \frac{1}{2}h^2 S\nabla^2 H(x)S\nabla H(x) + \mathcal{O}(h^2))\\
&+\frac{1}{2}D_2^2\overline{\nabla}H(x,x)(hS\nabla H(x)+\mathcal{O}(h^2), hS\nabla H(x)+\mathcal{O}(h^2)) + \mathcal{O}(h^3)\\
=& f + hSD_2\overline{\nabla}Hf + \frac{1}{2}h^2 SD_2\overline{\nabla}Hf'f + \frac{1}{2}h^2 SD_2^2\overline{\nabla}H(f,f) + \mathcal{O}(h^3),
\end{aligned}$$

where we have suppressed the argument $x$ of $f$, $D_2\overline{\nabla}H$ and $D_2^2\overline{\nabla}H$ in the last line. Furthermore, we use that

$$Q(x, x+\gamma hf(x)) = Q(x,x) + \gamma hD_2Q(x,x)(f,\cdot) + \mathcal{O}(h^2)$$

and (2.9) to get

$$\begin{aligned}
&SQ(x,x+\gamma hf(x))S\overline{\nabla}H(x,\hat{x})\\
&\quad= SQ(x,x)S\overline{\nabla}H(x,\hat{x}) + \gamma hSD_2Q(x,x)(f, S\overline{\nabla}H(x,\hat{x})) + \mathcal{O}(h^2)\\
&\quad= SQ(x,x)S(\nabla H(x) + D_2\overline{\nabla}H(x,x)(hS\nabla H(x)+\mathcal{O}(h^2))\\
&\qquad+\gamma hSD_2Q(x,x)(f, S(\nabla H(x)+\mathcal{O}(h)) + \mathcal{O}(h^2)\\
&\quad= SQ(x,x)f + hSQ(x,x)SD_2\overline{\nabla}H(x,x)f + \gamma hSD_2Q(x,x)(f,f) + \mathcal{O}(h^2)\\
&\quad= \frac{1}{2}f'f - SD_2\overline{\nabla}Hf + \frac{1}{2}hf'SD_2\overline{\nabla}Hf - hSD_2\overline{\nabla}HSD_2\overline{\nabla}Hf\\
&\qquad+\frac{1}{6}\gamma hf''(f,f) - \frac{1}{2}\gamma hSD_2^2\overline{\nabla}H(f,f) + \mathcal{O}(h^2),
\end{aligned}$$

where again we suppress the argument $x$ in the last expression. Thus the discrete gradient scheme (1.9) is of order 3 if $\overline{\nabla}H \in C^2(\mathbb{R}^d\times\mathbb{R}^d, \mathbb{R}^d)$ and $\overline{S}(x,\hat{x},h) = \overline{S}(x,h)$ is given by

$$\overline{S}(x,h) = S + hSQ(x, x+\frac{2}{3}hf(x))S + h^2S\big(Q(x,x)SQ(x,x) - \frac{1}{12}\nabla^2H(x)S\nabla^2H(x)\big)S.$$

Finding an approximation of $S$ that guarantees higher order of the discrete gradient method quickly becomes significantly more complicated, and results in increasingly complicated expressions for $\overline{S}(x,\hat{x},h)$. For example, it can be shown that one fourth order scheme of the form (1.9) is given by any $\overline{\nabla}H \in C^3(\mathbb{R}^d\times\mathbb{R}^d, \mathbb{R}^d)$ and

$$\begin{aligned}
\overline{S}(x,h) =& S + hS\big(\frac{8}{9}Q(x,z_3) + \frac{1}{9}Q(x,x)\big)S\\
&+h^2S\big(Q(x,z_2)SQ(x,z_2) - \frac{1}{12}\nabla^2H(z_1)S\nabla^2H(z_1)\big)S\\
&+h^3S\big(Q(x,x)SQ(x,x)SQ(x,x)\\
&-\frac{1}{12}\nabla^2H(x)S\nabla^2H(x)SQ(x,x) - \frac{1}{12}Q(x,x)S\nabla^2H(x)S\nabla^2H(x)\big)S,
\end{aligned} \tag{2.17}$$

where

$$z_1 = x + \frac{1}{2}hf(x), \qquad z_2 = x + \frac{2}{3}hf(x), \qquad z_3 = x + \frac{3}{4}hf(z_1).$$

Note that if we choose a symmetric discrete gradient, we have by Lemma 2.1 that $Q(x,x) = 0$, and many of the terms in (2.17) disappear. If we use the AVF discrete gradient, (2.17) simplifies to

$$\overline{S}(x,h) = S - \frac{1}{12}h^2 S\nabla^2 H(z_1)S\nabla^2 H(z_1)S. \tag{2.18}$$

This is very similar to the higher order AVF methods of Quispel and McLaren, as given in [26], applied to (1.4) with $S$ constant: if we replace $z_1$ in (2.18) by $x$, we get their third order scheme; if we replace $z_1$ by $\frac{x+\hat{x}}{2}$, we get their symmetric fourth order scheme.

Seeing as (2.17) simplifies considerably when the AVF discrete gradient is chosen, and since we in this case get a B-series method, we begin our generalization to arbitrary order by studying this case specifically in the chapter to follow.

# 3 A generalization of the AVF method

Let us recall the concept of B-series. Referring to the definitions in [13, Section III.1], we let $T$ be the set of rooted trees, built recursively from starting with $\tau = \bullet$ and letting $\tau = [\tau_1, \dots, \tau_m]$ be the tree obtained by grafting the roots of the trees $\tau_1, \dots, \tau_m$ to a new root. Furthermore, $F(\tau)$ is the elementary differential associated with the tree $\tau$, defined by $F(\bullet)(x) = f(x)$ and

$$F(\tau)(x) = f^{(m)}(x)\big(F(\tau_1)(x), \dots, F(\tau_m)(x)\big),$$

and $\sigma(\tau)$ is the symmetry coefficient for $\tau$, defined by $\sigma(\bullet) = 1$ and

$$\sigma(\tau) = \sigma(\tau_1)\cdots\sigma(\tau_m)\cdot \mu_1!\mu_2!\cdots, \tag{3.1}$$

where the integers $\mu_1, \mu_2, \dots$ count equal trees among $\tau_1, \dots, \tau_m$. Then, if $\phi : T \cup \{\emptyset\} \to \mathbb{R}$ is an arbitrary map, a *B-series* is a formal series

$$B(\phi, x) = \phi(\emptyset)x + \sum_{\tau \in T} \frac{h^{|\tau|}}{\sigma(\tau)}\phi(\tau)F(\tau)(x). \tag{3.2}$$

The exact solution of (1.1) can be written as the B-series $B(\frac{1}{\gamma}, x)$, where the coefficient $\gamma$ satisfies $\gamma(\emptyset) = \gamma(\bullet) = 1$ and

$$\gamma(\tau) = |\tau|\gamma(\tau_1)\cdots\gamma(\tau_m) \qquad \text{for } \tau = [\tau_1, \dots, \tau_m], \tag{3.3}$$

where $|\tau|$ is the order, i.e. the number of nodes, of $\tau$.

**Definition 3.1.** The *generalized AVF method* is given by

$$\frac{\hat{x} - x}{h} = \left(I + \sum_{n=2}^{p-1} h^n \sum_j b_{nj}\left(\prod_{k=1}^{n} f'(z_{njk}) + (-1)^n \prod_{k=1}^{n} f'(z_{nj(n-k+1)})\right)\right)\int_0^1 f((1-\xi)x + \xi\hat{x})\,\mathrm{d}\xi, \tag{3.4}$$

where each $z_{njk} := z_{njk}(x, \hat{x}, h) = B(\phi_{njk}, x)$ can be written as a B-series with $\phi(\emptyset) = 1$.

Note that we may alternatively write (3.4) in the slightly more compact form

$$\frac{\hat{x}-x}{h} = \sum_{n=0}^{p-1} h^n \sum_j b_{nj} \left( \prod_{k=1}^{n} f'(z_{njk}) + (-1)^n \prod_{k=1}^{n} f'(z_{nj(n-k+1)}) \right) \int_0^1 f((1-\xi)x + \xi\hat{x})\,\mathrm{d}\xi$$

with $\sum_j b_{0j} = \frac{1}{2}$.

**Theorem 3.1.** *When applied to* (1.1) *with* $f(x) = S\nabla H(x)$*, where* $S$ *is a constant skew-symmetric matrix, the scheme* (3.4) *preserves* $H$*, in that* $H(\hat{x}) = H(x)$*.*

*Proof.* With $f(x) = S\nabla H(x)$, (3.4) becomes

$$\frac{\hat{x}-x}{h} = \overline{S}(x,\hat{x},h)\overline{\nabla}_{\mathrm{AVF}} H(x,\hat{x}),$$

with

$$\overline{S}(x,\hat{x},h) = S + \sum_{n=2}^{p-1} h^n \sum_j b_{nj} \left( \prod_{k=1}^{n} S\nabla^2 H(z_{njk}) + (-1)^n \prod_{k=1}^{n} S\nabla^2 H(z_{nj(n-k+1)}) \right) S.$$

We have

$$\begin{aligned}
&\left( \prod_{k=1}^{n} S\nabla^2 H(z_{njk}) \cdot S + (-1)^n \prod_{k=1}^{n} S\nabla^2 H(z_{nj(n-k+1)}) \cdot S \right)^T \\
&\quad = S^T \prod_{k=1}^{n} \left( \nabla^2 H(z_{nj(n-k+1)})^T S^T \right) + (-1)^n S^T \prod_{k=1}^{n} \left( \nabla^2 H(z_{njk})^T S^T \right) \\
&\quad = (-1)^{i+1} S \prod_{k=1}^{n} \nabla^2 H(z_{nj(n-k+1)}) S - S \prod_{k=1}^{n} \nabla^2 H(z_{njk}) S \\
&\quad = -\left( \prod_{k=1}^{n} S\nabla^2 H(z_{njk}) \cdot S + (-1)^n \prod_{k=1}^{n} S\nabla^2 H(z_{nj(n-k+1)}) \cdot S \right),
\end{aligned}$$

and thus $\overline{S}(x,\hat{x},h)$ is a skew-symmetric matrix. □

Before considering the order conditions of the generalized AVF method, let us recall a couple of results from the literature on B-series.

**Lemma 3.2** ([13, Lemma III.1.9])**.** *Let* $B(a,x)$ *be a B-series with* $a(\emptyset) = 1$*. Then* $hf(B(a,x)) = B(a',x)$ *is also a B-series, with* $a'(\emptyset) = 0$*,* $a'(\bullet) = 1$ *and otherwise*

$$a'(\tau) = a(\tau_1)\cdots a(\tau_m) \quad \textit{for } \tau = [\tau_1,\ldots,\tau_m].$$

**Lemma 3.3** ([23, Theorem 2.2])**.** *Let* $B(a,x)$ *and* $B(b,x)$ *be two B-series with* $a(\emptyset) = 1$ *and* $b(\emptyset) = 0$*. Then* $hf'(B(a,x))B(b,x) = B(a\times b,x)$*, i.e. a B-series, with* $(a\times b)(\emptyset) = (a\times b)(\bullet) = 0$ *and otherwise*

$$(a\times b)(\tau) = \sum_{i=1}^{m} \prod_{j=1, j\neq i}^{m} a(\tau_j) b(\tau_i) \quad \textit{for } \tau = [\tau_1,\ldots,\tau_m].$$

Proposition 1 in [2] states that the standard AVF method is a B-series method. We build on the proof of that proposition to prove the following result.

**Proposition 3.4.** *The generalized AVF method* (3.4) *is a B-series method.*

*Proof.* First we define $\hat{e}: T \cup \{\emptyset\} \to \mathbb{R}$ by $\hat{e}(\emptyset) = 1$ and $\hat{e}(\tau) = 0$ for all $\tau \neq \emptyset$. Then, assuming that the solution $\hat{x}$ of (3.4) can be written as the B-series $\hat{x} = B(\Phi, x)$, we find the B-series

$$\begin{aligned} h\int_0^1 f\big((1-\xi)x + \xi\hat{x}\big)\,\mathrm{d}\xi &= h\int_0^1 f\big(B((1-\xi)\hat{e} + \xi\Phi, x)\big)\,\mathrm{d}\xi \\ &= \int_0^1 B\big(((1-\xi)\hat{e} + \xi\Phi)', x\big)\,\mathrm{d}\xi = B\Big(\int_0^1 ((1-\xi)\hat{e} + \xi\Phi)'\,\mathrm{d}\xi, x\Big). \end{aligned}$$

Setting $\theta := \int_0^1 ((1-\xi)\hat{e} + \xi\Phi)'\,\mathrm{d}\xi = \int_0^1 ((1-\xi)\hat{e})'\,\mathrm{d}\xi + \int_0^1 (\xi\Phi)'\,\mathrm{d}\xi = \int_0^1 (\xi\Phi)'\,\mathrm{d}\xi$, we get

$$\theta(\emptyset) = 0, \quad \theta(\bullet) = 1, \quad \theta([\tau_1, \ldots, \tau_m]) = \frac{1}{m+1}\Phi(\tau_1)\cdots\Phi(\tau_m). \tag{3.5}$$

Then we may rewrite (3.4) as

$$\begin{aligned} \hat{x} &= x + \left(I + \sum_{n=2}^{p-1} h^n \sum_j b_{nj}\left(\prod_{k=1}^n f'(B(\phi_{njk}, x)) + (-1)^n \prod_{k=1}^n f'(B(\phi_{nj(n-k+1)}, x))\right)\right) B(\theta, x) \\ &= x + B(\theta, x) + \sum_{n=2}^{p-1}\sum_j b_{nj}\big(B(\phi_{nj1} \times \cdots \times \phi_{njn} \times \theta, x) + (-1)^n B(\phi_{njn} \times \cdots \times \phi_{nj1} \times \theta, x)\big) \\ &= B(\Phi, x), \end{aligned}$$

with

$$\Phi = \hat{e} + \theta + \sum_{n=2}^{p-1}\sum_j b_{nj}\big(\phi_{nj1} \times \cdots \times \phi_{njn} \times \theta + (-1)^n \phi_{njn} \times \cdots \times \phi_{nj1} \times \theta\big). \tag{3.6}$$

□

Comparing the B-series of the exact solution and the B-series of the solution of (3.4), and noting that the elementary differentials are independent, we immediately get the following result.

**Theorem 3.5.** *The generalized AVF method* (3.4) *is of order $p$ if and only if*

$$\Phi(\tau) = \frac{1}{\gamma(\tau)} \quad \textit{for } |\tau| \leq p, \tag{3.7}$$

*where $\Phi$ is given by* (3.6) *and $\gamma$ is given by* (3.3).

The terms $\Phi(\tau)$ can be found from (3.6) by applying Lemma 3.3 recursively, as illustrated by the following example.

| $\lvert\tau\rvert$ | $F(\tau)^i$ | $\tau$ | $\sigma(\tau)$ | $\gamma(\tau)$ | $\Phi(\tau)$ |
|---|---|---|---|---|---|
| 1 | $f^i$ | $\bullet$ | 1 | 1 | 1 |
| 2 | $f^i_j f^j$ | | 1 | 2 | $\frac{1}{2}$ |
| 3 | $f^i_{jk} f^j f^k$ | | 2 | 3 | $\frac{1}{3}$ |
| | $f^i_j f^j_k f^k$ | | 1 | 6 | $\frac{1}{4} + 2\sum_j b_{2j}$ |
| 4 | $f^i_{jkl} f^j f^k f^l$ | | 6 | 4 | $\frac{1}{4}$ |
| | $f^i_{jk} f^j f^k_l f^l$ | | 1 | 8 | $\frac{1}{6} + \sum_{j,k} b_{2j}\phi_{2jk}(\bullet)$ |
| | $f^i_j f^j_{kl} f^k f^l$ | | 2 | 12 | $\frac{1}{6} + 2\sum_{j,k} b_{2j}\phi_{2j1}(\bullet)$ |
| | $f^i_j f^j_k f^k_l f^l$ | | 1 | 24 | $\frac{1}{8} + 2\sum_j b_{2j}$ |

Table 1: Elementary differentials and their coefficients in the B-series of the solution of (3.4), up to fourth order.

**Example 3.1.** Consider $\tau =$ , and assume we have found $\Phi$ for all trees up to and including order four already, as given in Table 1. We have

$$\theta(\;) = \frac{1}{3}\Phi(\bullet)\Phi(\;) = \frac{1}{3}(\frac{1}{4} + 2\sum_j b_{2j}) = \frac{1}{12} + \frac{2}{3}\sum_j b_{2j}.$$

Then we calculate

$$\begin{aligned}(\phi_{2j1} \times \phi_{2j2} \times \theta)(\;) &= \phi_{2j1}(\bullet)(\phi_{2j2} \times \theta)(\;) + \phi_{2j1}(\;)(\phi_{2j2} \times \theta)(\bullet) \\ &= \phi_{2j1}(\bullet)(\phi_{2j2} \times \theta)(\;) = \phi_{2j1}(\bullet)\phi_{2j2}(\emptyset)\theta(\;) = \frac{1}{2}\phi_{2j1}(\bullet),\end{aligned}$$

where we have used in the second equality that $(\phi_{2j2} \times \theta)(\bullet) = \phi_{2j2}(\emptyset)\theta(\emptyset) = 0$. Similarly we find $(\phi_{2j2} \times \phi_{2j1} \times \theta)(\;) = \frac{1}{2}\phi_{2j2}(\bullet)$. Furthermore,

$$\begin{aligned}(\phi_{3j1} \times \phi_{3j2} \times \phi_{3j3} \times \theta)(\;) &= \phi_{3j1}(\bullet)(\phi_{3j2} \times \phi_{3j3} \times \theta)(\;) = \phi_{3j1}(\bullet)\phi_{3j2}(\emptyset)(\phi_{3j3} \times \theta)(\;) \\ &= \phi_{3j1}(\bullet)\phi_{3j3}(\emptyset)\theta(\bullet) = \phi_{3j1}(\bullet),\end{aligned}$$

and $(\phi_{3j3} \times \phi_{3j2} \times \phi_{3j1} \times \theta)(\;) = \phi_{3j3}(\bullet)$. Hence,

$$\Phi(\;) = \frac{1}{12} + \frac{2}{3}\sum_j b_{2j} + \frac{1}{2}\sum_j b_{2j}(\phi_{2j1}(\bullet) + \phi_{2j2}(\bullet)) + \sum_j b_{3j}(\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)).$$

Now, if we assume the order condition (3.7) to be satisfied for all trees up to and including order four, we can replace $\sum_j b_{2j} = -\frac{1}{24}$ and $\sum_j b_{2j}(\phi_{2j1}(\bullet) + \phi_{2j2}(\bullet)) = -\frac{1}{24}$ in the above expression, use that $\gamma(\,)= 30$, and get that (3.7) is satisfied for  if and only if

$$\sum_j b_{3j}(\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)) = -\frac{1}{720}. \tag{3.8}$$

## 3.1 Construction of higher order schemes

As the size of the trees grows, finding $\Phi(\tau)$ from (3.6) can become quite a cumbersome operation. Furthermore, we observe from Table 1 that there are some equivalent order conditions for different trees. Before presenting more convenient techniques for finding order conditions for the generalized AVF method, let us define some more concepts related to B-series and trees.

First, recall that the Butcher product of two trees $u = [u_1, \dots, u_m]$ and $v = [v_1, \dots, v_n]$ is given by $u \circ v = [u_1, u_2, \dots, u_m, v]$. This operation is neither associative nor commutative, and in contrast to the practice in [13], we here take the product of several factors without parentheses to mean evaluation from right to left:

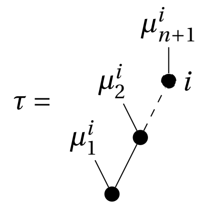
$$u_1 \circ u_2 \circ \cdots \circ u_k := u \circ (u_2 \circ (\cdots \circ u_k)).$$

Given a forest $\mu = (\tau_1, \dots, \tau_m)$, the tree obtained by grafting the roots of every tree in $\mu$ to a new root is denoted by $[\mu] = [\tau_1, \dots, \tau_m]$. Moreover, $\mu^{-1}(\tau)$ denotes the forest such that $[\mu^{-1}(\tau)] = \tau$. We extend the maps $\phi : T \cup \{\emptyset\} \to \mathbb{R}$ and $\gamma : T \cup \{\emptyset\} \to \mathbb{R}$ to forests by the letting $\phi(\mu) = \prod_{i=1}^m \phi(\tau_i)$ and $\gamma(\mu) = \prod_{i=1}^m \gamma(\tau_i)$ for $\mu = (\tau_1, \dots, \tau_m)$.

Consider now a tree $\tau$ consisting of $|\tau|$ nodes. We may number every tree from 1 to $|\tau|$, starting at the root and going from left to right on the increasing levels above. For a given node $i \in [1, \dots, |\tau|]$ on level $n+1$, there exists a unique set of forests $\hat{\tau}^i = \{\mu_1^i, \dots, \mu_{n+1}^i\}$ such that

$$\tau = [\mu_1^i] \circ [\mu_2^i] \circ \cdots \circ [\mu_{n+1}^i].$$

That is, labeling node $i$,




**Proposition 3.6.** *The* $\Phi$ *of* (3.7) *can alternatively be found by*

$$\Phi(\tau) = \hat{e}(\tau) + \theta(\tau) + \sum_{i \text{ s.t. } n \geq 2} \Lambda(\hat{\tau}^i) \tag{3.9}$$

*where* $\hat{e}(\emptyset) = 1$ *and* $\hat{e}(\tau) = 0$ *for all* $\tau \neq \emptyset$, $\theta(\emptyset) = 0$, $\theta(\bullet) = 1$,

$$\theta([\tau_1, \dots, \tau_m]) = \frac{1}{m+1}\Phi(\tau_1)\cdots\Phi(\tau_m),$$

*and*

$$\Lambda(\hat{\tau}^i) = \theta([\mu^i_{n+1}])\sum_j b_{nj}\left(\phi_{nj1}(\mu^i_1)\cdots\phi_{njn}(\mu^i_n) + (-1)^n\phi_{njn}(\mu^i_1)\cdots\phi_{nj1}(\mu^i_n)\right). \tag{3.10}$$

*Proof.* Define $n_i$ so that $n_i+1$ is the level of node $i$. Collect the children of node $i$ in the set $C_i$. We have

$$[\mu^i_{n_i+1}] = [\mu^k_{n_k}]\circ[\mu^k_{n_k+1}] \qquad \text{for all } k\in C_i,$$

and thus

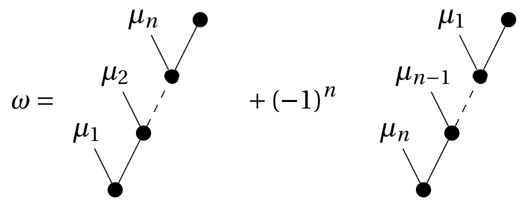
$$(a\times b)([\mu^i_{n_i+1}]) = \sum_{k\in C_i} a(\mu^k_{n_k})b([\mu^k_{n_k+1}]).$$

Note also that $\mu^i_{n_i} = \mu^k_{n_i} = \mu^k_{n_k-1}$ if $k\in C_i$. Then we get

$$\begin{aligned}
(\phi_{nj1}\times\cdots\times\phi_{njn}\times\theta)(\tau) &= (\phi_{nj1}\times\cdots\times\phi_{njn}\times\theta)([\mu^1_1])\\
&= \sum_{i_1\in C_1}\phi_{nj1}(\mu^{i_1}_1)(\phi_{nj2}\times\cdots\times\phi_{njn}\times\theta)([\mu^{i_1}_2])\\
&= \sum_{i_1\in C_1}\phi_{nj1}(\mu^{i_1}_1)\sum_{i_2\in C_{i_1}}\phi_{nj2}(\mu^{i_2}_2)(\phi_{nj3}\times\cdots\times\phi_{njn}\times\theta)([\mu^{i_2}_3])\\
&= \sum_{i_1\in C_1}\sum_{i_2\in C_{i_1}}\phi_{nj1}(\mu^{i_2}_1)\phi_{nj2}(\mu^{i_2}_2)(\phi_{nj3}\times\cdots\times\phi_{njn}\times\theta)([\mu^{i_2}_3])\\
&\;\;\vdots\\
&= \sum_{i_1\in C_1}\sum_{i_2\in C_{i_1}}\cdots\sum_{i_n\in C_{i_{n-1}}}\phi_{nj1}(\mu^{i_n}_1)\cdots\phi_{njn}(\mu^{i_n}_n)\theta([\mu^{i_n}_{n+1}])\\
&= \sum_{i \text{ on level } n+1}\phi_{nj1}(\mu^i_1)\cdots\phi_{njn}(\mu^i_n)\theta([\mu^i_{n+1}]).
\end{aligned}$$

Inserting this and the corresponding result for $(\phi_{njn}\times\cdots\times\phi_{nj1}\times\theta)(\tau)$ in (3.6), we get (3.10). □

In [8, 3], conditions are derived for a B-series method to be energy preserving when applied to the system (1.6). In [26], while giving the AVF method as one such method, Quispel and McLaren present a general form of what they call energy-preserving linear combinations of rooted trees:




Here we give their result as a lemma, which is proved later by the proof of the more general Theorem 4.7.

**Lemma 3.7.** *Let $\mu_1,\ldots,\mu_n$ be $n$ arbitrary forests. Then, if $f(x) = S\nabla H(x)$ for some skew-symmetric constant matrix $S$, we have that $F(\omega)(x)\cdot\nabla H(x) = 0$ for*

$$\omega = [\mu_1]\circ[\mu_2]\circ\cdots[\mu_n]\circ[\emptyset] + (-1)^n\,[\mu_n]\circ[\mu_{n-1}]\circ\cdots[\mu_1]\circ[\emptyset]. \tag{3.11}$$

There is a connection between (3.9) and Lemma 3.7 such that instead of order conditions for every tree, we can calculate order conditions for every energy-preserving linear combination. To see this we start by collecting the leaf nodes, i.e. nodes with no children, of the tree $\tau$ in a set $I_l$ and the other nodes in the set $I_n$. If node $i \in I_n$, we may then use the relation

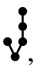
$$\Lambda(\{\mu_1^i, \dots, \mu_n^i, \mu_{n+1}^i\}) = \theta([\mu_{n+1}^i])\Lambda(\{\mu_1^i, \dots, \mu_n^i, \emptyset\})$$

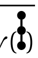
to find $\Lambda(\hat{\tau}^i)$ from the previously calculated $\Lambda$ for a smaller tree. Then if lower order conditions are satisfied, we have numerical values for these $\Lambda$. The leaf nodes on the other hand, with their corresponding $\hat{\tau}^i = \{\mu_1^i, \dots, \mu_n^i, \emptyset\}$, gives an energy-preserving linear combination (3.11) which $\tau$ belongs to. If $i$ is on level two, this combination is simply $\tau - \tau = 0$, and accordingly $\Lambda$ is not calculated for these nodes in (3.9). Moreover, leaves on the same level have identical $\hat{\tau}^i$. Thus, a tree with leaves on $m$ different levels above level two will belong to at most $m$ non-zero energy-preserving linear combinations (3.11).

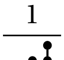
If we assume the conditions for order $< p$ to be satisfied, we may replace (3.7) by

$$\sum_{i \in I_l} \Lambda(\hat{\tau}^i) = \frac{1}{\gamma(\tau)} - \hat{e}(\tau) - \sum_{i \in I_n} \frac{\Lambda(\{\mu_1^i, \dots, \mu_n^i, \emptyset\})}{(|\mu_{n+1}^i| + 1)\gamma(\mu_{n+1}^i)}, \tag{3.12}$$

where $|\mu|$ denotes the number of trees in the forest $\mu$. Note that $\Lambda(\{\emptyset\}) = 1$ and hence $\Lambda(\hat{\tau}^1) = \theta(\tau)$. Then we can calculate the numerical value for the right hand side and, if $\tau$ has leaves on only one level $> 2$, find an order condition for both $\tau$ and the other tree in the combination (3.11). This warrants an example.

**Example 3.2.** Consider again the tree $\tau =$ , which is part of the energy-preserving linear combination $\omega =$  $-$ . Ignoring the two nodes on level 2, there are three nodes to calculate $\Lambda$ for: $i = 1$, $i = 4$ and $i = 5$. We find

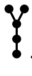
$\Lambda(\hat{\tau}^1) = \dfrac{1}{(2+1)\gamma(\bullet)\gamma(\text{tree})} = \dfrac{1}{3 \cdot 1 \cdot 6} = \dfrac{1}{18}$, with tree 

$\Lambda(\hat{\tau}^4) = \dfrac{1}{(1+1)\gamma(\bullet)}\Lambda(\{\bullet, \emptyset, \emptyset\}) = \dfrac{1}{2\gamma(\bullet)}\left(\dfrac{1}{\gamma(\text{tree})} - \dfrac{1}{3\gamma(\bullet)\gamma(\text{tree})}\right) = \dfrac{1}{2}\left(\dfrac{1}{8} - \dfrac{1}{6}\right) = -\dfrac{1}{48}$, with trees  and 

$$\Lambda(\hat{\tau}^5) = \Lambda(\{\bullet, \emptyset, \emptyset, \emptyset\}) = \sum_j b_{3j}\big(\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)\big).$$

The right hand side of (3.12) becomes

$$\frac{1}{\gamma(\tau)} - \frac{1}{18} - (-\frac{1}{48}) = \frac{1}{30} - \frac{1}{18} + \frac{1}{48} = -\frac{1}{720},$$

and we have the order condition (3.8) for the linear combination  $-$ .

If there are leaves on $r > 1$ different levels levels above level two, things get slightly more complicated. Then we get $r$ different terms on the left hand side of (3.12) and we need to consider the order condition for $\tau$ and the $r$ trees it forms energy-preserving linear combinations with, so that we get an equation for every energy-preserving combination of these trees, also those not including $\tau$. This is illustrated by the following example.

**Example 3.3.** The tree forms energy-preserving combinations with both and . Thus we have to calculate (3.12) for all three trees to find order conditions for the corresponding linear combinations. Starting with $\tau =$ , which has three nodes above level two, two leaves and one non-leaf, we get

$$\Lambda(\hat{\tau}^4) = \Lambda(\{\bullet, \;, \emptyset\}) = \sum_j b_{2j}\big(\phi_{2j1}(\bullet)\phi_{2j2}(\;) + \phi_{2j2}(\bullet)\phi_{2j1}(\;)\big),$$

$$\Lambda(\hat{\tau}^5) = \frac{1}{(1+1)\gamma(\bullet)}\Lambda(\{\bullet, \bullet, \emptyset\}) = \frac{1}{2\gamma(\bullet)}\frac{1}{2}\Big(\frac{1}{\gamma(\;)} - \frac{1}{3\gamma(\bullet)\gamma(\;)}\Big) = \frac{1}{2}\frac{1}{2}\Big(\frac{1}{15} - \frac{1}{9}\Big) = -\frac{1}{90},$$

$$\Lambda(\hat{\tau}^6) = \Lambda(\{\bullet, \bullet, \emptyset, \emptyset\}) = \sum_j b_{3j}\big(\phi_{3j1}(\bullet)\phi_{3j2}(\bullet) - \phi_{3j3}(\bullet)\phi_{3j2}(\bullet)\big)$$
$$= \sum_j b_{3j}\phi_{3j2}(\bullet)(\phi_{3j1} - \phi_{3j3})(\bullet).$$

For the right hand side of (3.12), we get

$$\frac{1}{\gamma(\tau)} - \frac{1}{(2+1)\gamma(\bullet)\gamma(\;)} - (-\frac{1}{90}) = \frac{1}{48} - \frac{1}{3\cdot 1\cdot 8} + \frac{1}{90} = -\frac{7}{720},$$

and hence the order condition for is

$$\sum_j b_{2j}\big(\phi_{2j1}(\bullet)\phi_{2j2}(\;) + \phi_{2j2}(\bullet)\phi_{2j1}(\;)\big) + \sum_j b_{3j}\phi_{3j2}(\bullet)(\phi_{3j1} - \phi_{3j3})(\bullet) = -\frac{7}{720}. \qquad (3.13)$$

Similarly we calculate (3.12) for ,

$$\sum_{jk} b_{2j}\phi_{2jk}(\;) - 2\sum_j b_{3j}\phi_{3j2}(\bullet)(\phi_{3j1} - \phi_{3j3})(\bullet) = -\frac{1}{120}, \qquad (3.14)$$

and for ,

$$\sum_{jk} b_{2j}\phi_{2jk}(\;) + 2\sum_j b_{2j}\big(\phi_{2j1}(\bullet)\phi_{2j2}(\;) + \phi_{2j2}(\bullet)\phi_{2j1}(\;)\big) = -\frac{1}{36}. \qquad (3.15)$$

Combining (3.13), (3.14) and (3.15), we get the equivalent system of equations

$$\sum_j b_{3j}\phi_{3j2}(\bullet)(\phi_{3j1}(\bullet)-\phi_{3j3}(\bullet)) = \frac{1}{240}+\alpha, \tag{3.16}$$

$$\sum_j b_{2j}(\phi_{2j1}(\bullet)\phi_{2j2}(\;)+\phi_{2j1}(\;)\phi_{2j2}(\bullet)) = -\frac{1}{72}-\alpha, \tag{3.17}$$

$$\sum_{j,k} b_{2j}\phi_{2jk}(\;) = 2\alpha, \tag{3.18}$$

where the choice of $\alpha \in \mathbb{R}$ is arbitrary. The order conditions (3.16)–(3.18) can be associated to the linear combinations $-$ , $+$ and $+$ , respectively.

By considering the order conditions in Table 2, we find a fifth order scheme of the form (3.4) given by

$$\begin{aligned}\frac{\hat{x}-x}{h} =& \Big(I-\frac{5}{136}h^2\big(f'(z_2)f'(z_3)+f'(z_3)f'(z_2)\big)-\frac{1}{102}h^2f'(x)f'(x)\\ &+\frac{1}{288}h^3\big(f'(x)f'(x)f'(z_1)+f'(z_1)f'(x)f'(x)\big)\\ &+\frac{1}{120}h^4f'(x)f'(x)f'(x)f'(x)\Big)\int_0^1 f((1-\xi)x+\xi\hat{x})\,\mathrm{d}\xi,\end{aligned} \tag{3.19}$$

where

$$z_1 = x+\frac{2}{5}hf(x), \qquad z_2 = x+\frac{17+\sqrt{17}}{30}hf(z_1), \qquad z_3 = x+\frac{17-\sqrt{17}}{30}hf(z_1).$$

A symmetric sixth order scheme is given by

$$\begin{aligned}\frac{\hat{x}-x}{h} =& \Big(I-\frac{13}{360}h^2f'\big(\bar{x}+\frac{\sqrt{13}}{26}hf(\bar{x}-\frac{3\sqrt{13}}{26}hf(\bar{x}))\big)f'\big(\bar{x}-\frac{\sqrt{13}}{26}hf(\bar{x}+\frac{3\sqrt{13}}{26}hf(\bar{x}))\big)\\ &-\frac{13}{360}h^2f'\big(\bar{x}-\frac{\sqrt{13}}{26}hf(\bar{x}+\frac{3\sqrt{13}}{26}hf(\bar{x}))\big)f'\big(\bar{x}+\frac{\sqrt{13}}{26}hf(\bar{x}-\frac{3\sqrt{13}}{26}hf(\bar{x}))\big)\\ &-\frac{1}{180}h^2f'(x)f'(x)-\frac{1}{180}h^2f'(\hat{x})f'(\hat{x})\\ &+\frac{1}{720}h^3f'(\bar{x}-\frac{1}{2}hf(\bar{x}))f'(\bar{x})f'(\bar{x}+\frac{1}{2}hf(\bar{x}))\\ &-\frac{1}{720}h^3f'(\bar{x}+\frac{1}{2}hf(\bar{x}))f'(\bar{x})f'(\bar{x}-\frac{1}{2}hf(\bar{x}))\\ &+\frac{1}{120}h^4f'(\bar{x})f'(\bar{x})f'(\bar{x})f'(\bar{x})\Big)\int_0^1 f((1-\xi)x+\xi\hat{x})\,\mathrm{d}\xi,\end{aligned} \tag{3.20}$$

where $\bar{x} = \frac{x+\hat{x}}{2}$. If we wish to calculate the matrix in front of the integral explicitly, we have a

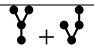
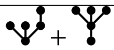
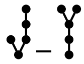
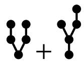
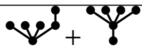
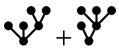
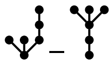
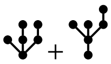
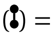
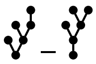
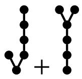
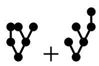
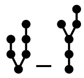
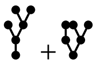
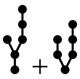
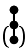
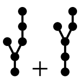
| $\lvert\tau\rvert$ | $\omega$ | Order condition |
|---|---|---|
| 1 | $\bullet$ | – |
| 2 | – | – |
| 3 |  | $\sum_j b_{2j} = -\frac{1}{24}$ |
| 4 |  | $\sum_{j,k} b_{2j}\phi_{2jk}(\bullet) = -\frac{1}{24}$ |
| 5 |  | $\sum_{j,k} b_{2j}\phi_{2jk}(\bullet)^2 = -\frac{1}{40}$ |
| |  | $\sum_j b_{2j}\phi_{2j1}(\bullet)\phi_{2j2}(\bullet) = -\frac{1}{90}$ |
| |  | $\sum_j b_{3j}(\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)) = -\frac{1}{720}$ |
| |  | $\sum_{j,k} b_{2j}\phi_{2jk}($$) = -\frac{1}{60}$ |
| |  | $\sum_j b_{4j} = \frac{1}{240}$ |
| 6 |  | $\sum_{j,k} b_{2j}\phi_{2jk}(\bullet)^3 = -\frac{1}{60}$ |
| |  | $\sum_j b_{2j}(\phi_{2j1}(\bullet)^2\phi_{2j2}(\bullet) + \phi_{2j1}(\bullet)\phi_{2j2}(\bullet)^2) = -\frac{1}{72}$ |
| |  | $\sum_j b_{3j}(\phi_{3j1}(\bullet)^2 - \phi_{3j3}(\bullet)^2) = -\frac{1}{720}$ |
| |  | $\sum_{j,k} b_{2j}\phi_{2jk}(\bullet)\phi_{2jk}($$) = -\frac{1}{96}$ |
| |  | $\sum_j b_{3j}\phi_{3j2}(\bullet)(\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)) = \frac{1}{240} + \alpha_1$ |
| |  | $\sum_j b_{4j}(\phi_{4j1}(\bullet) + \phi_{4j4}(\bullet)) = \frac{1}{240}$ |
| |  | $\sum_j b_{2j}(\phi_{2j1}(\bullet)\phi_{2j2}($$) + \phi_{2j1}($$)\phi_{2j2}(\bullet)) = -\frac{1}{72} - \alpha_1$ |
| |  | $\sum_j b_{3j}(\phi_{3j1}($$) - \phi_{3j3}($$)) = -\frac{1}{180} - \alpha_2$ |
| |  | $\sum_{j,k} b_{2j}\phi_{2jk}($$) = 2\alpha_1$ |
| |  | $\sum_{j,k} b_{2j}\phi_{2jk}($$) = \alpha_2$ |
| |  | $\sum_j b_{4j}(\phi_{4j2}(\bullet) + \phi_{4j3}(\bullet)) = -\frac{1}{1440} - \alpha_2$ |

Table 2: Energy-preserving linear combinations of elementary differentials, and their associated order conditions for the scheme (3.4), up to sixth order. The coefficients $\alpha_1, \alpha_2 \in \mathbb{R}$ are arbitrary.

non-symmetric sixth order scheme given by

$$\begin{aligned}\frac{\hat{x}-x}{h} = \Big(I - \frac{13}{360}h^2\big(f'(z_6)f'(z_7) + f'(z_7)f'(z_6)\big) - \frac{1}{180}h^2\big(f'(x)f'(x) + f'(z_1)f'(z_1)\big) \\ + \frac{1}{720}h^3\big(f'(x)f'(z_2)f'(z_3) - f'(z_3)f'(z_2)f'(x)\big) \\ + \frac{1}{120}h^4 f'(z_2)f'(z_2)f'(z_2)f'(z_2)\Big)\int_0^1 f((1-\xi)x + \xi\hat{x})\,\mathrm{d}\xi,\end{aligned} \tag{3.21}$$

with

$$z_1 = x + \frac{1}{4}hf(x) + \frac{3}{4}hf\big(x + \frac{2}{3}hf(x + \frac{1}{3}hf(x))\big), \qquad z_2 = x + \frac{1}{2}hf(x), \qquad z_3 = x + hf(z_2),$$
$$z_4 = \frac{1}{2}(x + z_3) - \frac{3\sqrt{13}}{26}hf(z_2), \qquad z_5 = \frac{1}{2}(x + z_3) + \frac{3\sqrt{13}}{26}hf(z_2),$$
$$z_6 = \frac{1}{2}(x + z_1) + \frac{\sqrt{13}}{26}hf(z_4), \qquad z_7 = \frac{1}{2}(x + z_1) - \frac{\sqrt{13}}{26}hf(z_5).$$

# 4 AVF discrete gradient methods for general skew-gradient systems

We will now build on the results of the previous paper by generalizing the results to the situation where $S(x)$ in the skew-gradient system (1.4) is not necessarily constant. Consider therefore now an ODE of the form (1.4), and set again $g := \nabla H$. By Taylor expansion of $x$ around $t = t_0$ we get

$$\begin{aligned} x(t_0+h) =& x + hSg + \frac{h^2}{2}(S'gSg + Sg'Sg) + \frac{h^3}{6}(S''g(Sg,Sg) + 2S'g'(Sg,Sg) + Sg''(Sg,Sg) \\ & + S'gS'gSg + S'gSg'Sg + Sg'S'gSg + Sg'Sg'Sg) + \mathcal{O}(h^4), \end{aligned}$$

where $x := x(t_0)$, and $S$, $g$ and their derivatives are evaluated in $x$. Introducing the notation $f^\circ := S'g$ and $f^\bullet := Sg'$, we can write this in the abbreviated form

$$\begin{aligned} x(t_0+h) =& x + hf + \frac{h^2}{2}(f^\circ f + f^\bullet f) + \frac{h^3}{6}(f^{\circ\circ}(f,f) + 2f^{\circ\bullet}(f,f) + f^{\bullet\bullet}(f,f) \\ & + f^\circ f^\circ f + f^\circ f^\bullet f + f^\bullet f^\circ f + f^\bullet f^\bullet f) + \mathcal{O}(h^4). \end{aligned} \tag{4.1}$$

## 4.1 Skew-gradient systems and P-series

A *P-series* is given by

$$P(\phi,(x,y)) = \begin{pmatrix} \phi(\emptyset)x + \sum_{\tau\in TP_\bullet} \frac{h^{|\tau|}}{\sigma(\tau)}\phi(\tau)F(\tau)(x,y) \\ \phi(\emptyset)y + \sum_{\tau\in TP_\circ} \frac{h^{|\tau|}}{\sigma(\tau)}\phi(\tau)F(\tau)(x,y) \end{pmatrix}, \tag{4.2}$$

where $TP$ is the set of rooted bi-colored trees and $TP_\bullet$ and $TP_\circ$ are the subsets of $TP$ whose roots are black and white, respectively [13, Section III.2]. The bi-colored trees are built recursively; starting with $\bullet$ and $\circ$, we let $\tau = [\tau_1,\dots,\tau_m]_\bullet$ be the tree you get by grafting the roots of $\tau_1,\dots,\tau_m$ to a black root and $\tau = [\tau_1,\dots,\tau_m]_\circ$ the tree you get by grafting $\tau_1,\dots,\tau_m$ to a white root. No subscript, i.e. $\tau = [\tau_1,\dots,\tau_m]$, means grafting to a black root.

The exact solution of a partitioned system

$$\begin{aligned} \dot{x} &= f(x,y), \quad x(t_0) = x_0, \\ \dot{y} &= g(x,y), \quad y(t_0) = y_0, \end{aligned} \tag{4.3}$$

can be written as $(x(t_0+h), y(t_0+h)) = P(1/\gamma,(x_0,y_0))$, where the coefficient $\gamma$ is given by $\gamma(\emptyset) = \gamma(\bullet) = \gamma(\circ) = 1$ and (3.3). As noted in [5], setting $f(x,y) := S(y)\nabla H(x)$, the skew-gradient system (1.4) can be written as (4.3) with $g = f$. When $g = f$, all coefficients and the elementary differentials $F(\tau)$ in (4.2) are given independent of the color of the root. Thus for the system (1.4), it suffices to consider

$$P(\phi, x) = \phi(\emptyset)x + \sum_{\tau \in TP_\bullet} \frac{h^{|\tau|}}{\sigma(\tau)} \phi(\tau) F(\tau)(x), \tag{4.4}$$

and we have that the exact solution of (1.4) can be written as $x(t_0 + h) = P(1/\gamma, x_0)$. Breaking slightly with convention, we define a P-series to be the single row version (4.4) in the remainder of this paper. Denoting black-rooted subtrees by $\tau_i$ and white-rooted subtrees by $\bar{\tau}_i$, the elementary differentials $F(\tau)$ for the skew-gradient system are given recursively by $F(\bullet)(x) = F(\circ)(x) = S(x)\nabla H(x)$, and

$$F(\tau)(x) = S^{(l)} D^m \nabla H(F(\tau_1)(x), \ldots, F(\tau_m)(x), F(\bar{\tau}_1)(x), \ldots, F(\bar{\tau}_l)(x)) \tag{4.5}$$

for both $\tau = [\tau_1, \ldots, \tau_m, \bar{\tau}_1, \ldots, \bar{\tau}_l]_\bullet$ and $\tau = [\tau_1, \ldots, \tau_m, \bar{\tau}_1, \ldots, \bar{\tau}_l]_\circ$. The bi-colored trees in $TP_\bullet$ and their corresponding elementary differentials $F$ are given up to order three in Table 3. The number of trees grows very quickly with the order; see `https://oeis.org/A000151`.

| $\lvert\tau\rvert$ | $F(\tau)^i$ | $F(\tau)$ | $\tau$ | $\alpha(\tau)$ | $\gamma(\tau)$ | $\sigma(\tau)$ |
|---|---|---|---|---|---|---|
| 1 | $S^i_j g^j$ | $f$ | $\bullet$ | 1 | 1 | 1 |
| 2 | $S^i_{jk} g^j S^k_l g^l$ | $f^\circ f$ |  | 1 | 2 | 1 |
| | $S^i_j g^j_k S^k_l g^l$ | $f^\bullet f$ |  | 1 | 2 | 1 |
| 3 | $S^i_{jkm} g^j S^k_l g^l S^m_n g^n$ | $f^{\circ\circ}(f,f)$ |  | 1 | 3 | 2 |
| | $S^i_{jk} g^j_m S^k_l g^l S^m_n g^n$ | $f^{\circ\bullet}(f,f)$ |  | 2 | 3 | 1 |
| | $S^i_j g^j_{km} S^k_l g^l S^m_n g^n$ | $f^{\bullet\bullet}(f,f)$ |  | 1 | 3 | 2 |
| | $S^i_{jk} g^j S^k_{lm} g^l S^m_n g^n$ | $f^\circ f^\circ f$ |  | 1 | 6 | 1 |
| | $S^i_j g^j_k S^k_{lm} g^l S^m_n g^n$ | $f^\bullet f^\circ f$ |  | 1 | 6 | 1 |
| | $S^i_{jk} g^j S^k_l g^l_m S^m_n g^n$ | $f^\circ f^\bullet f$ |  | 1 | 6 | 1 |
| | $S^i_j g^j_k S^k_l g^l_m S^m_n g^n$ | $f^\bullet f^\bullet f$ |  | 1 | 6 | 1 |

Table 3: Bi-colored trees and their elementary differentials up to third order.

The following lemma is Lemma III.2.2 in [13] amended to fit our setting.

**Lemma 4.1.** *Let $P(a, x)$ and $P(b, x)$ be two P-series with $a(\emptyset) = b(\emptyset) = 1$. Then*

$$hS(P(a, x))\nabla H(P(b, x)) = P(a \vee b, x),$$

*where $(a \vee b)(\emptyset) = 0$, $(a \vee b)(\bullet) = 1$, and*

$$(a \vee b)(\tau) = a(\tau_1) \cdots a(\tau_m) b(\bar{\tau}_1) \cdots b(\bar{\tau}_l) \quad \textit{for } \tau = [\tau_1, \ldots, \tau_m, \bar{\tau}_1, \ldots, \bar{\tau}_l]_\bullet.$$

**Proposition 4.2.** *The AVF discrete gradient scheme*

$$\frac{\hat{x}-x}{h} = S\Big(\frac{x+\hat{x}}{2}\Big)\int_0^1 \nabla H((1-\xi)x+\xi\hat{x})\,\mathrm{d}\xi \tag{4.6}$$

*is a second order P-series method.*

*Proof.* As in the proof of Proposition 3.4, we define $\hat{e}$ by $\hat{e}(\emptyset)=1$ and $\hat{e}(\tau)=0$ for all $\tau\neq\emptyset$. Now, assume that the solution $\hat{x}$ of (4.6) can be written as the P-series $\hat{x}=P(\Phi,x)$. Then, using Lemma 4.1, we find the P-series

$$\begin{aligned}
hS\Big(\frac{x+\hat{x}}{2}\Big)\int_0^1 \nabla H((1-\xi)x+\xi\hat{x})\,\mathrm{d}\xi &= hS\big(P\big(\frac{1}{2}\hat{e}+\frac{1}{2}\Phi,x\big)\big)\int_0^1 \nabla H\big(P((1-\xi)\hat{e}+\xi\Phi,\,x)\big)\,\mathrm{d}\xi\\
&= \int_0^1 hS\big(P\big(\frac{1}{2}\hat{e}+\frac{1}{2}\Phi,\,x\big)\big)\nabla H\big(P((1-\xi)\hat{e}+\xi\Phi,\,x)\big)\,\mathrm{d}\xi\\
&= P\Big(\int_0^1 \big(\big(\frac{1}{2}\hat{e}+\frac{1}{2}\Phi\big)\vee\big((1-\xi)\hat{e}+\xi\Phi\big)\big)\,\mathrm{d}\xi,\,x\Big).
\end{aligned}$$

Thus we get $\Phi=\hat{e}+\int_0^1\big(\big(\frac{1}{2}\hat{e}+\frac{1}{2}\Phi\big)\vee\big((1-\xi)\hat{e}+\xi\Phi\big)\big)\,\mathrm{d}\xi=\hat{e}+\int_0^1\big(\big(\frac{1}{2}\Phi\big)\vee\big(\xi\Phi\big)\big)\,\mathrm{d}\xi$. That is,

$$\Phi(\emptyset)=1,\quad \Phi(\bullet)=1,\quad \Phi([\tau_1,\dots,\tau_m,\bar{\tau}_1,\dots,\bar{\tau}_l])=\frac{1}{(m+1)2^l}\Phi(\tau_1)\cdots\Phi(\tau_m)\Phi(\bar{\tau}_1)\cdots\Phi(\bar{\tau}_l).$$

Writing out the first few terms of the series, we have

$$\begin{aligned}
\hat{x}=x+hf+\frac{h^2}{2}(f^{\circ}f+f^{\bullet}f)+h^3\big(&\frac{1}{8}f^{\circ\circ}(f,f)+\frac{1}{4}f^{\circ\bullet}(f,f)+\frac{1}{6}f^{\bullet\bullet}(f,f)\\
&+\frac{1}{4}f^{\circ}f^{\circ}f+\frac{1}{4}f^{\circ}f^{\bullet}f+\frac{1}{4}f^{\bullet}f^{\circ}f+\frac{1}{4}f^{\bullet}f^{\bullet}f\big)+\mathcal{O}(h^4),
\end{aligned}$$

which, after comparing with the expanded exact solution (4.1), we see is of order two. □

The following lemma is obtained in a manner similar to Lemma 3.3, i.e. Theorem 2.2 in [23], and hence we present it without its proof.

**Lemma 4.3.** *Let $P(a,x)$, $P(b,x)$ and $P(c,x)$ be three P-series with $a(\emptyset)=b(\emptyset)=1$ and $c(\emptyset)=0$. Then*

$$hS(P(a,x))\nabla^2 H(P(b,x))P(c,x)=P((a,b)\times c,x)$$

*with $((a,b)\times c)(\emptyset)=((a,b)\times c)(\bullet)=0$ and otherwise*

$$((a,b)\times c)(\tau)=\sum_{i=1}^{m}\prod_{j=1,j\neq i}^{m}\prod_{k=1}^{l}a(\bar{\tau}_k)b(\tau_j)c(\tau_i)\quad \textit{for } \tau=[\tau_1,\dots,\tau_m,\bar{\tau}_1,\dots,\bar{\tau}_l]. \tag{4.7}$$

Note that $\{\emptyset\}$ counts as both a black-rooted and a white-rooted tree. Hence we have e.g.

$$((a,b)\times c)([\bullet,\circ])=a(\circ)b(\emptyset)c(\bullet)=a(\bullet)c(\bullet),$$

where we also use that $a(\circ) = a(\bullet)$.

We now present a subclass of the AVF discrete gradient method, for which we will find order conditions using Lemma 4.1 and Lemma 4.3. This subclass is every AVF discrete gradient method for which the approximation of $S(x)$ can be written on the form

$$\begin{aligned}\overline{S}(x,\hat{x},h) = \sum_{n=0}^{p-1} h^n \sum_j b_{nj}\Big(&\prod_{k=1}^{n} S(\bar{z}_{njk})\nabla^2 H(z_{njk})\cdot S(\bar{z}_{nj(n+1)})\\ &+(-1)^n S(\bar{z}_{nj(n+1)})\prod_{k=1}^{n}\nabla^2 H(z_{nj(n-k+1)})S(\bar{z}_{nj(n-k+1)})\Big),\end{aligned} \tag{4.8}$$

where, if $\hat{x}$ is the solution of

$$\frac{\hat{x}-x}{h} = \overline{S}(x,\hat{x},h)\overline{\nabla}_{\mathrm{AVF}}H(x,\hat{x}),$$

each $z_{njk} := z_{njk}(x,\hat{x},h) = P(\phi_{njk},x)$ and each $\bar{z}_{njk} := \bar{z}_{njk}(x,\hat{x},h) = P(\psi_{njk},x)$ can be written as a P-series with $\phi_{njk}(\emptyset) = \psi_{njk}(\emptyset) = 1$ for all $n,j,k$. We require that $\sum_j b_{0j} = \frac{1}{2}$, which ensures that (4.8) is a consistent approximation of $S(x)$.

**Theorem 4.4.** *The discrete gradient scheme* (1.9) *with the AVF discrete gradient* (1.10) *and the approximation of* $S(x)$ *given by* (4.8) *is a P-series method.*

*Proof.* Generalizing the argument in the proof of Proposition 4.2, we find the P-series

$$h\,S\big(P(a,x)\big)\int_0^1 \nabla H((1-\xi)x+\xi\hat{x})\,\mathrm{d}\xi = P\Big(\int_0^1 \big(a\vee\big((1-\xi)\hat{e}+\xi\Phi\big)\big)\,\mathrm{d}\xi,\,x\Big),$$

where $\bar{\theta}(a) := \int_0^1 \big(a\vee\big((1-\xi)\hat{e}+\xi\Phi\big)\big)\,\mathrm{d}\xi = \int_0^1 \big(a\vee\xi\Phi\big)\big)\,\mathrm{d}\xi$, so that $\bar{\theta}(a)(\emptyset) = 0$, $\bar{\theta}(a)(\bullet) = 1$, and

$$\bar{\theta}(a)([\tau_1,\ldots,\tau_m,\bar{\tau}_1,\ldots,\bar{\tau}_l]) = \frac{1}{m+1}\Phi(\tau_1)\cdots\Phi(\tau_m)a(\bar{\tau}_1)\cdots a(\bar{\tau}_l). \tag{4.9}$$

Thus we may write the solution $\hat{x}$ found from applying the scheme (1.9) with the AVF discrete gradient (1.10) and $\overline{S}(x,\hat{x},h)$ given by (4.8) as

$$\begin{aligned}\hat{x} &= x + \sum_{n=0}^{p-1} h^n \sum_j b_{nj}\left(\prod_{k=1}^{n} S(P(\psi_{njk},x))\nabla^2 H(P(\phi_{njk},x))\cdot P(\bar{\theta}(\psi_{nj(n+1)}),x)\right.\\ &\qquad\left. +(-1)^n \prod_{k=1}^{n} S(P(\psi_{nj(n-k+2)},x))\nabla^2 H(P(\phi_{nj(n-k+1)},x))\cdot P(\bar{\theta}(\psi_{nj1}),x)\right)\\ &= x + \sum_{n=0}^{p-1}\sum_j b_{nj}\big(P((\psi_{nj1},\phi_{nj1})\times(\psi_{nj2},\phi_{nj2})\times\cdots\times(\psi_{njn},\phi_{njn})\times\bar{\theta}(\psi_{nj(n+1)}),x)\\ &\qquad +(-1)^n P((\psi_{nj(n+1)},\phi_{njn})\times(\psi_{njn},\phi_{nj(n-1)})\times\cdots\times(\psi_{nj2},\phi_{nj1})\times\bar{\theta}(\psi_{nj1}),x)\big)\\ &= P(\Phi,x),\end{aligned} \tag{4.10}$$

with

$$\begin{aligned}\Phi =&\hat{e}+\sum_{n=0}^{p-1}\sum_{j} b_{nj}\big((\psi_{nj1},\phi_{nj1})\times\cdots\times(\psi_{njn},\phi_{njn})\times\bar{\theta}(\psi_{nj(n+1)})\\ &+(-1)^{n}(\psi_{nj(n+1)},\phi_{njn})\times\cdots\times(\psi_{nj2},\phi_{nj1})\times\bar{\theta}(\psi_{nj1})\big).\end{aligned} \tag{4.11}$$

□

**Theorem 4.5.** *The AVF discrete gradient method with $\overline{S}$ given by* (4.8) *is of order $p$ if and only if*

$$\Phi(\tau)=\frac{1}{\gamma(\tau)}\quad \textit{for } |\tau|\le p. \tag{4.12}$$

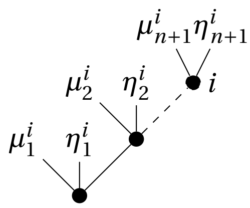
The values $\Phi(\tau)$ can be found from (4.11) using (4.7) recursively and then (4.9). However, a more convenient approach is derived in the next section.

## 4.2 Order conditions

This section is devoted to generalization of the results in Section 3.1 to the cases where $S(x)$ is not necessarily constant. To that end, for a tree $\tau\in TP_{\bullet}$, we cut off all branches between black and white nodes and denote the mono-colored tree we are left with by $\tau^{b}$. We number the nodes in that tree as before, from 1 to $|\tau^{b}|$, and reattach the cut-off parts to the tree to get $\tau$ again. Let $\mu$ denote a forest of black-rooted trees and $\eta$ a forest of white-rooted trees. Then, for a given node $i\in[1,\ldots,|\tau^{b}|]$ on level $n+1$, there exists a unique set of forests $\hat{\tau}^{i}=\{(\mu_1^i,\eta_1^i),\ldots,(\mu_{n+1}^i,\eta_{n+1}^i)\}$ such that

$$\tau=[(\mu_1^i,\eta_1^i)]\circ[(\mu_2^i,\eta_2^i)]\circ\cdots\circ[(\mu_{n+1}^i,\eta_{n+1}^i)].$$

That is,

$\tau =$ 


Now we can generalize Proposition 3.6 as follows.

**Proposition 4.6.** *The $\Phi$ of* (4.11) *can be found by*

$$\Phi(\tau)=\hat{e}(\tau)+\sum_{i=1}^{|\tau^{b}|}\Lambda(\hat{\tau}^{i}) \tag{4.13}$$

*where $\hat{e}(\emptyset)=1$ and $\hat{e}(\tau)=0$ for all $\tau\neq\emptyset$, and*

$$\begin{aligned}\Lambda(\hat{\tau}^{i}) =&\theta([\mu_{n+1}^{i}])\sum_{j} b_{nj}\Big(\psi_{nj1}(\eta_1^i)\phi_{nj1}(\mu_1^i)\cdots\psi_{njn}(\eta_n^i)\phi_{njn}(\mu_n^i)\psi_{nj(n+1)}(\eta_{n+1}^i)\\ &+(-1)^{n}\psi_{nj(n+1)}(\eta_1^i)\phi_{njn}(\mu_1^i)\psi_{njn}(\eta_2^i)\cdots\phi_{nj1}(\mu_n^i)\psi_{nj1}(\eta_{n+1}^i)\Big),\end{aligned} \tag{4.14}$$

*with*

$$\theta([\tau_1,\ldots,\tau_m])=\frac{1}{m+1}\Phi(\tau_1)\cdots\Phi(\tau_m).$$

*Proof.* Defining $n_i$ and $C_i$ as in the proof of Proposition 3.6, we have

$$[(\mu^i_{n_i+1},\eta^i_{n_i+1})] = [(\mu^k_{n_k},\eta^k_{n_k})]\circ[(\mu^k_{n_k+1},\eta^k_{n_k+1})] \quad \text{for all } k\in C_i,$$
$$((a,b)\times c)([[(\mu^i_{n_i+1},\eta^i_{n_i+1})]]) = \sum_{k\in C_i} a(\eta^k_{n_k})b(\mu^k_{n_k})c([\mu^k_{n_k+1},\eta^k_{n_k+1}]).$$

Observe that $\bar\theta(a)([\mu,\eta]) = a(\eta)\theta([\mu])$. For $n=0$ we have

$$\bar\theta(\psi_{0j1})(\tau) = \bar\theta(\psi_{0j1})([\mu^1_1,\eta^1_1]) = \psi_{0j1}(\eta^1_1)\theta([\mu^1_1]),$$

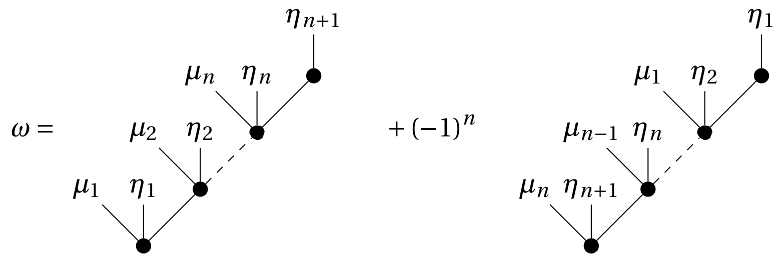
and for $n>0$ we get

$$\begin{aligned}
&((\psi_{nj1},\phi_{nj1})\times\cdots\times(\psi_{njn},\phi_{njn})\times\bar\theta(\psi_{nj(n+1)}))(\tau)\\
&\quad= ((\psi_{nj1},\phi_{nj1})\times\cdots\times(\psi_{njn},\phi_{njn})\times\bar\theta(\psi_{nj(n+1)}))([\mu^1_1,\eta^1_1])\\
&\quad= \sum_{i_1\in C_1}\psi_{nj1}(\eta^{i_1}_1)\phi_{nj1}(\mu^{i_1}_1)((\psi_{nj2},\phi_{nj2})\times\cdots\times(\psi_{njn},\phi_{njn})\times\bar\theta(\psi_{nj(n+1)}))([\mu^{i_1}_2,\eta^{i_1}_2])\\
&\quad\ \vdots\\
&\quad= \sum_{i_1\in C_1}\cdots\sum_{i_n\in C_{i_{n-1}}}\psi_{nj1}(\eta^{i_n}_1)\phi_{nj1}(\mu^{i_n}_1)\cdots\psi_{njn}(\eta^{i_n}_n)\phi_{njn}(\mu^{i_n}_n)\bar\theta(\psi_{nj(n+1)})([\mu^{i_n}_{n+1},\eta^{i_n}_{n+1}])\\
&\quad= \sum_{i\text{ on level } n+1}\psi_{nj1}(\eta^i_1)\phi_{nj1}(\mu^i_1)\cdots\psi_{njn}(\eta^i_n)\phi_{njn}(\mu^i_n)\psi_{nj(n+1)}(\eta^i_{n+1})\theta([\mu^i_{n+1}]).
\end{aligned}$$

Inserting this and the corresponding result for $((\psi_{nj(n+1)},\phi_{njn})\times\cdots\times(\psi_{nj2},\phi_{nj1})\times\bar\theta(\psi_{nj1}))(\tau)$ in (4.11), we get (4.14). $\square$

Note that if $\tau$ only has black nodes, we have $\Lambda(\hat\tau^1) = \theta(\tau)\sum_j b_{0j}(\psi_{0j1}(\emptyset)+\psi_{0j1}(\emptyset)) = \theta(\tau)$, and also $\Lambda(\hat\tau^i) = 0$ for all nodes $i$ on level 2. Thus (4.13) simplifies to (3.9).

Like for the constant $S$ case, the order conditions can be given for energy-preserving linear combinations of elementary differentials instead for each elementary differential. In the following generalization of Lemma 3.7, we state that the energy-preserving linear combinations of bi-colored rooted trees are given by




**Theorem 4.7.** *Let $\mu_1,\mu_2,\dots,\mu_n$ be arbitrary forests of black-rooted trees and $\eta_1,\eta_2,\dots,\eta_{n+1}$ arbitrary forests of white-rooted trees. Given $f(x) = S(x)\nabla H(x)$, where $S(x)$ is a skew-symmetric matrix, and elementary differentials defined by* (4.5)*, the linear combinations of trees given by*

$$\omega = [(\mu_1,\eta_1)]\circ\cdots\circ[(\mu_n,\eta_n)]\circ[\eta_{n+1}] + (-1)^n[(\mu_n,\eta_{n+1})]\circ\cdots\circ[(\mu_1,\eta_2)]\circ[\eta_1] \tag{4.15}$$

*are energy-preserving in the sense that $F(\omega)(x)\cdot\nabla H(x) = 0$.*

*Proof.* For any forest of black-rooted trees $\mu_j$, we have $F([\mu_j]\circ[\emptyset]) = SB_jS\nabla H$ for some symmetric matrix $B_j$, suppressing the argument $x$. Similarly, for a forest of white-rooted trees $\eta_j$, we have $F([\eta_j]) = W_j\nabla H$ for some skew-symmetric matrix $W_j$. Note that the empty forest is considered both a black-rooted and a white-rooted forest, and accordingly we have $F([\emptyset]\circ[\emptyset]) = F(\,) = S(\nabla^2 H)S\nabla H$ and $F([\emptyset]) = F(\bullet) = S\nabla H$. For these matrices $B_j$ and $W_j$ corresponding to the forests $\mu_j$ and $\eta_j$, we get

$$F\big([(\mu_1,\eta_1)]\circ\cdots\circ[(\mu_n,\eta_n)]\circ[\eta_{n+1}]\big) = W_1B_1W_2B_2\cdots B_nW_{n+1}\nabla H.$$

We have

$$(W_1B_1W_2B_2\cdots B_nW_{n+1})^T = \begin{cases} -W_{n+1}B_nW_nB_{n-1}\cdots B_1W_1 & \text{if } n \text{ even,} \\ W_{n+1}B_nW_nB_{n-1}\cdots B_1W_1 & \text{if } n \text{ odd.} \end{cases}$$

Thus $F(\omega)(x)$ is a skew-symmetric matrix times $\nabla H(x)$, and the statement in the above theorem follows directly. $\square$

**Example 4.1.** We show that the combination $+$ is energy-preserving.

$$: \quad (f^\bullet f^{\bullet\bullet}(f,f^\circ f))^i = S^i_j g^j_k S^k_l g^l_{mo} S^m_n g^n S^o_{pq} g^p S^q_r g^r = S^i_j g^j_k S^k_l g^l_{mo} S^m_n g^n S^o_{pq} S^q_r g^r g^p.$$

$$: \quad (f^{\circ\bullet\bullet}(f,f,f^\bullet f))^i = S^i_{jk} g^j_{mo} S^k_l g^l S^m_n g^n S^o_p g^p_q S^q_r g^r = S^i_{jk} S^k_l g^l g^j_{mo} S^m_n g^n S^o_r g^r_q S^q_p g^p.$$

For this linear combination on the form (4.15), we have $\eta_1 = \eta_2 = \emptyset, \eta_3 = \circ$, $\mu_1 = \emptyset, \mu_2 = \bullet$, with the corresponding matrices $W_1 = W_2 = S, (W_3)^i_j = S^i_{jk}S^k_l g^l$ and $B_1 = \nabla^2 H, (B_2)^j_m = g^j_{km}S^k_l g^l$. Thus we get

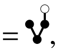
$$+ = f^\bullet f^{\bullet\bullet}(f,f^\circ f) + f^{\circ\bullet\bullet}(f,f,f^\bullet f) = Z\nabla H,$$

where $Z \coloneqq S(\nabla^2 H)SB_2W_3 + W_3B_2S(\nabla^2 H)S$ is a skew-symmetric matrix.

For bi-colored trees, we define a node on the tree $\tau$ to be a leaf if it is a leaf on the corresponding cut tree $\tau^b$ by the definition of leaves given in the previous chapter. We let $I_l$ be the set of leaves and $I_n$ the set of non-leaf nodes which are also in $\tau^b$, so that $I_l \cup I_n = [1,\dots,|\tau^b|]$. In contrast to the case with mono-colored trees, a leaf $i$ on level one or two of a bi-colored tree may give rise to a non-zero energy-preserving linear combination; it does so if and only if $\eta^i_k \neq \emptyset$ for any $k = 1,2$. Accordingly, $\Lambda(\hat{\tau}^i)$ is calculated in (4.13) also when $n = 0,1$. Furthermore, two leaves $i$ and $j$ on the same level will belong to two different energy-preserving combinations if $\eta^i_{n+1} \neq \eta^j_{n+1}$. Therefore we now simply state that a tree with $r$ leaves, also including the lower two levels, belong to at most $r$ non-zero linear combinations. We thus get $r$ terms on the left hand side of

$$\sum_{i\in I_l}\Lambda(\hat{\tau}^i) = \frac{1}{\gamma(\tau)} - \hat{e}(\tau) - \sum_{i\in I_n}\frac{\Lambda(\{(\mu^i_1,\eta^i_1),\dots,(\mu^i_n,\eta^i_n),(\emptyset,\eta^i_{n+1})\})}{(|\mu^i_{n+1}|+1)\gamma(\mu^i_{n+1})}, \tag{4.16}$$

which is equivalent to (4.12) if we assume the conditions for lower order to be satisfied.

**Example 4.2.** Consider the tree $\tau =$ , which is part of the energy-preserving linear combination  $-$ . Assume that the order conditions up to and including order three are all satisfied. The cut tree $\tau^b =$  has three nodes of which two are leaves. Node number 2 is a leaf on level 2 with $\eta_1^2 = \eta_2^2 = \emptyset$, and thus gives $\Lambda(\hat{\tau}^2) = 0$. We find for the other two,

$\Lambda(\hat{\tau}^1) = \Lambda(\{((\bullet,$  $),\emptyset)\}) = \dfrac{1}{(|\mu_1^1|+1)\gamma(\mu_1^1)}\Lambda\big(\{(\emptyset,\emptyset)\}\big) = \dfrac{1}{(2+1)\gamma(\bullet)\gamma(}$  $)}\dfrac{1}{\gamma(\bullet)} = \dfrac{1}{6},$

$$\Lambda(\hat{\tau}^3) = \Lambda(\{(\bullet,\emptyset),(\emptyset,\circ)\}) = \sum_j b_{1j}(\phi_{1j1}(\bullet)\psi_{1j2}(\circ) - \psi_{1j1}(\circ)\phi_{1j1}(\bullet)) = \sum_j b_{1j}\phi_{1j1}(\bullet)(\psi_{1j2} - \psi_{1j1})(\bullet).$$

For the right hand side of (4.16) we get

$$\frac{1}{\gamma(\tau)} - \Lambda(\hat{\tau}^1) = \frac{1}{8} - \frac{1}{6} = -\frac{1}{24},$$

and thus the order condition

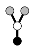
$$\sum_j b_{1j}\phi_{1j1}(\bullet)(\psi_{1j2} - \psi_{1j1})(\bullet) = -\frac{1}{24}$$

for the energy-preserving linear combination  $-$ .

Even though the number of black-rooted bi-colored trees grows very quickly, e.g. to 26 for $|\tau| = 4$ and 107 for $|\tau| = 5$, finding and satisfying the order conditions is not as daunting a task as it might first appear. First of all, it suffices to find order conditions for the non-zero linear combinations given by (4.15). Moreover, a couple key observations simplifies the process further:

- The large number of trees $\tau$ for which $\tau^b = \bullet$, i.e. trees with no black nodes on level 2, are all energy-preserving. They can be written $\tau = [\eta_1^1]$, and their order condition is given by

$$2\sum_j b_{0j}\psi_{0j1}(\eta_1^1) = \frac{1}{\gamma(\tau)}.$$

- For trees that are identical except for the colors of the descendants of white nodes, it suffices to calculate one order condition. E.g. for  we have the order condition $2b_{0j}\psi_{0j1}($$) = \frac{1}{12}$, where each of the gray nodes may be black or white. To satisfy these conditions, it is natural to require that $\bar{z}_{0j1}$ in (4.8) is a B-series *up to* order $p-1$.

From the order conditions displayed in Table 4 we find that one second order scheme is given by (1.9) using the AVF discrete gradient (1.10) and an explicit skew-symmetric approximation of $S$ given by $\overline{S}(x,\cdot,h) = S(x + \frac{1}{2}hf(x))$. A third order scheme is obtained if we instead use the skew-symmetric approximation of $S$ explicitly given by

$$\begin{aligned}\overline{S}(x,\cdot,h) = &\frac{1}{4}S(x) + \frac{3}{4}S(z_2) + \frac{1}{4}h\big(S(z_1)\nabla^2H(x)S(x) - S(x)\nabla^2H(x)S(z_1)\big)\\ &-\frac{1}{12}h^2\,S(x)\nabla^2H(x)S(x)\nabla^2H(x)S(x),\end{aligned} \tag{4.17}$$

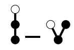
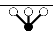
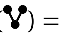
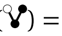
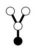
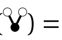
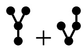
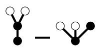
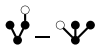
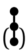
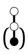
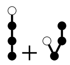
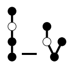
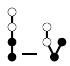
| $\lvert\tau\rvert$ | $\omega$ | Order condition |
|---|---|---|
| 1 | $\bullet$ | $2\sum_j b_{0j} = 1$ |
| 2 |  | $2\sum_j b_{0j}\psi_{0j1}(\bullet) = \frac{1}{2}$ |
| 3 |  | $2\sum_j b_{0j}\psi_{0j1}(\bullet)^2 = \frac{1}{3}$ |
| |  | $\sum_j b_{2j} = -\frac{1}{24}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{6}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{6}$ |
| |  | $\sum_j b_{1j}(\psi_{1j2} - \psi_{1j1})(\bullet) = -\frac{1}{12}$ |
| 4 |  | $2\sum_j b_{0j}\psi_{0j1}(\bullet)^3 = \frac{1}{4}$ |
| |  | $2\sum_j b_{2j}\psi_{2j2}(\bullet) = -\frac{1}{24}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{12}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{12}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{12}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}(\bullet)\psi_{0j1}($$) = \frac{1}{8}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}(\bullet)\psi_{0j1}($$) = \frac{1}{8}$ |
| |  | $\sum_j b_{2j}(\phi_{2j1} + \phi_{2j2})(\bullet) = -\frac{1}{24}$ |
| |  | $\sum_j b_{1j}(\psi_{1j2}(\bullet)^2 - \psi_{1j1}(\bullet)^2) = -\frac{1}{12}$ |
| |  | $\sum_j b_{1j}\phi_{1j1}(\bullet)(\psi_{1j1} - \psi_{1j0})(\bullet) = -\frac{1}{24}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{24}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{24}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{24}$ |
| |  | $2\sum_j b_{0j}\psi_{0j1}($$) = \frac{1}{24}$ |
| |  | $\sum_j b_{2j}(\psi_{2j1} + \psi_{2j3})(\bullet) = -\frac{1}{24}$ |
| |  | $\sum_j b_{1j}(\psi_{1j2} - \psi_{1j1})($$) = -\frac{1}{24}$ |
| |  | $\sum_j b_{1j}(\psi_{1j2} - \psi_{1j1})($$) = -\frac{1}{24}$ |

Table 4: Linear combinations $\omega$ of bi-colored black-rooted trees corresponding to energy-preserving elementary differentials of $f(x) = S(x)\nabla H(x)$, where $S(x)$ is a skew-symmetric matrix, as well as their associated order conditions for the discrete gradient method (1.9) with the AVF discrete gradient (1.10) and $\overline{S}(x,\hat{x},h)$ given by (4.8).

where $z_1 = x + \frac{1}{3}hf(x)$, $z_2 = x + \frac{2}{3}hf(z_1)$.

A symmetric fourth order scheme is given by (1.9) using the AVF discrete gradient (1.10) and the skew-symmetric approximation of $S$

$$\begin{aligned}\overline{S}(x,\hat{x},h) = &\frac{1}{2}S\Big(\bar{x} - \frac{1}{\sqrt{12}}hf\big(\bar{x} + \frac{1}{\sqrt{12}}hf(\bar{x})\big)\Big) + \frac{1}{2}S\Big(\bar{x} + \frac{1}{\sqrt{12}}hf\big(\bar{x} - \frac{1}{\sqrt{12}}hf(\bar{x})\big)\Big)\\ &+ \frac{1}{2}hS\big(\bar{x} + \frac{1}{12}hf(\bar{x})\big)\nabla^2 H(\bar{x})S\big(\bar{x} - \frac{1}{12}hf(\bar{x})\big)\\ &- \frac{1}{2}hS\big(\bar{x} - \frac{1}{12}hf(\bar{x})\big)\nabla^2 H(\bar{x})S\big(\bar{x} + \frac{1}{12}hf(\bar{x})\big)\\ &- \frac{1}{12}h^2 S(\bar{x})\nabla^2 H(\bar{x})S(\bar{x})\nabla^2 H(\bar{x})S(\bar{x}),\end{aligned} \tag{4.18}$$

where $\bar{x} = \frac{x+\hat{x}}{2}$. Another fourth order scheme is obtained if we instead use the explicit skew-symmetric approximation of $S$ found by

$$\begin{aligned}\overline{S}(x,\cdot,h) = &\frac{1}{2}(S(z_5+z_6) + S(z_5-z_6)) + \frac{1}{12}h\left(S(z_2)\nabla^2 H(z_1)S(x) - S(x)\nabla^2 H(z_1)S(z_2)\right)\\ &- \frac{1}{12}h^2 S(z_1)\nabla^2 H(z_1)S(z_1)\nabla^2 H(z_1)S(z_1),\end{aligned} \tag{4.19}$$

where

$$\begin{aligned} z_1 &= x + \frac{1}{2}hf(x), & z_3 &= x + hf(z_2), & z_5 &= \frac{1}{3}(x + z_1 + z_2) + \frac{1}{12}(-z_3 + z_4),\\ z_2 &= x + hf(z_1), & z_4 &= x + hf(z_3), & z_6 &= \frac{\sqrt{3}}{36}(7x - 2z_1 - 4z_2 + z_3 - 2z_4).\end{aligned}$$

# 5 Order conditions for general discrete gradient methods

We will now generalize the results of the two previous chapters to discrete gradient methods with a general discrete gradient, as defined by (1.7)–(1.8). To that end, we introduce two new series in the vein of B- and P-series, as well as related tree structures.

## 5.1 The constant $S$ case

Consider mono-colored rooted trees whose nodes can have two different shapes: the circle shape of the nodes in trees of B-series, but also a triangle shape. Let $TG$ be the set of such trees whose leaves are always circles. That is, from the first tree $\bullet$, every tree $\tau \in TG$ can be built recursively through

$$[\tau_1,\dots,\tau_m]_\bullet, \quad [\tau_1,\dots,\tau_m]_\blacktriangle, \quad \tau_1,\dots,\tau_m \in TG,$$

which denotes the grafting of the trees $\tau_1,\dots,\tau_m$ to a root $\bullet$ or $\blacktriangle$, respectively. The elementary differentials $F(\tau)$ corresponding to a tree $\tau \in TG$ are likewise defined recursively by $F(\bullet)(x) = f(x) = S\nabla H(x)$ and

$$F(\tau)(x) = \begin{cases} SD^m\nabla H(x)(F(\tau_1)(x),\dots,F(\tau_m)(x)) & \text{for } \tau = [\tau_1,\dots,\tau_m]_\bullet,\\ SD_2^{m-1}Q(x,x)(F(\tau_1)(x),\dots,F(\tau_m)(x)) & \text{for } \tau = [\tau_1,\dots,\tau_m]_\blacktriangle.\end{cases}$$

We can then define a generalization of B-series which includes these elementary differentials.

**Definition 5.1.** A G-series is a formal series of the form

$$G(\phi, x) = \phi(\emptyset)x + \sum_{\tau\in TG} \frac{h^{|\tau|}}{\sigma(\tau)}\phi(\tau)F(\tau)(x), \tag{5.1}$$

where $\phi : TG\cup\{\emptyset\} \to \mathbb{R}$ is an arbitrary mapping, and the symmetry coefficient $\sigma$ is given by (3.1).

The G-series of the exact solution is given by $x(t_0 + h) = G(\xi, x(t_0))$, with

$$\xi(\tau) = \begin{cases} \frac{1}{\gamma(\tau)} & \text{if } \tau \in T, \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

For use in the remainder of this paper, we generalize the Butcher product by the definition

$$u \circ v = [u_1, \dots, u_m, v]_\star, \quad \text{for } u = [u_1, \dots, u_m]_\star, \quad \star \in \{\bullet, \blacktriangle\}.$$

Furthermore, we let $|\tau|$ denote the total number of nodes in $\tau$, and $|\tau|_\star$ the number of nodes of type $\star$. Let $SG$ be the set of tall trees in $TG$; that is, the set of threes with only one node on each level. For a tree $\tau \in TG$, number every tree from 1 to $|\tau|$, as before. For any node $i$ on level $n+1$, we define the stem $s^i \in SG$ to be the tall tree consisting of the nodes connecting the root to node $i$, including the root and node $i$. Denote the $j^{\text{th}}$ node of $s^i$ by $s^i_j$, so that $s^i_1$ is the root and $s^i_{n+1} = i$. Then we have a unique set of forests $\hat{\tau}^i = \{\mu^i_1, \dots, \mu^i_{n+1}\}$ such that

$$\tau = [\mu^i_1]_{s^i_1} \circ [\mu^i_2]_{s^i_2} \circ \dots \circ [\mu^i_{n+1}]_{s^i_{n+1}}.$$

That is,

$$\tau = \begin{array}{ccc} & & \mu^i_{n+1} \\ & & \mid \\ & \mu^i_2 & s^i_{n+1} \\ & \backslash & / \\ \mu^i_1 & s^i_2 & \\ \backslash & / & \\ s^i_1 & & \end{array}$$

The following lemma is a generalization of Lemma 3.3 to G-series. Its proof is very similar to the proof of [23, Theorem 2.2], and hence omitted.

**Lemma 5.1.** *Let $G(a, x)$ and $G(b, x)$ be two G-series with $a(\emptyset) = 1$ and $b(\emptyset) = 0$. Then the G-series $hS\nabla^2 H(G(a, x))G(b, x) = G(a \times b, x)$ is given by $(a \times b)(\emptyset) = (a \times b)(\bullet) = 0$ and otherwise*

$$(a \times b)(\tau) = \begin{cases} \sum_{i=1}^m \prod_{j=1, j\neq i}^m a(\tau_j)b(\tau_i) & \textit{for } \tau = [\tau_1, \dots, \tau_m]_\bullet, \\ 0 & \textit{for } \tau = [\tau_1, \dots, \tau_m]_\blacktriangle. \end{cases}$$

*Moreover, $hSQ(x, G(a, x))G(b, x) = G(a \otimes b, x)$, with $(a \otimes b)(\emptyset) = (a \otimes b)(\bullet) = 0$ and otherwise*

$$(a \otimes b)(\tau) = \begin{cases} 0 & \textit{for } \tau = [\tau_1, \dots, \tau_m]_\bullet, \\ \sum_{i=1}^m \prod_{j=1, j\neq i}^m a(\tau_j)b(\tau_i) & \textit{for } \tau = [\tau_1, \dots, \tau_m]_\blacktriangle. \end{cases}$$

To every stem $s \in SG$ of height $n+1 = |s|$, we associate coefficients $b_{sj}$ and $\phi_{sjk}$. Letting $s_k$ be the $k^{\text{th}}$ node of $s$, we define the function

$$R(\phi_{sjk}, x) := \begin{cases} \nabla^2 H(G(\phi_{sjk}, x)) & \text{if } s_k = \bullet, \\ Q(x, G(\phi_{sjk}, x)) & \text{if } s_k = \blacktriangle. \end{cases}$$

Then we have $hSR(\phi_{sjk}, x)G(b, x) = G(\phi_{sjk} \diamond b)$, with $(\phi_{sjk} \diamond b)(\emptyset) = (\phi_{sjk} \diamond b)(\bullet) = 0$ and

$$(\phi_{sjk} \diamond b)(\tau) = \begin{cases} \sum_{i=1}^m \prod_{j=1, j\neq i}^m \phi_{sjk}(\tau_j) b(\tau_i) & \text{for } \tau = [\tau_1, \ldots, \tau_m]_{s_k}, \\ 0 & \text{if root of } \tau \neq s_k. \end{cases}$$

Consider now the class of skew-symmetric and consistent approximations to $S$ that can be written on the form

$$\overline{S}(x, y, h) = \sum_{s \in SG} h^n \sum_j b_{sj} \left( \prod_{k=1}^n SR(\phi_{sjk}, x) + (-1)^{|s|_\bullet - 1} \prod_{k=1}^n SR(\phi_{sj(n-k+1)}, x) \right) S \tag{5.3}$$

whenever $y$ is the solution of

$$\frac{y - x}{h} = \overline{S}(x, y, h) \overline{\nabla} H(x, y),$$

with $\phi_{sjk}(\emptyset) = 1$ for every $s, j, k$, and with $\sum_j b_{\bullet j} = \frac{1}{2}$.

**Lemma 5.2.** *The discrete gradient method* (1.9) *with* $\overline{S}(x, \hat{x}, h)$ *given by* (5.3) *and* $\overline{\nabla} H \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ *is a G-series method when applied to a constant S skew-gradient system* (1.6).

*Proof.* Assume that the solution $\hat{x}$ of (1.9) with $\overline{S}(x, \hat{x}, h)$ given by (5.3) can be written as the G-series $\hat{x} = G(\Phi, x)$. Then, using Lemma 2.2 and $\overline{\nabla} H(x, x) = \nabla H(x)$,

$$\begin{aligned} hS\overline{\nabla} H(x, \hat{x}) &= hS \sum_{m=0}^\infty \frac{1}{m!} D_2^m \overline{\nabla} H(x, x)(G(\Phi, x) - x)^m \\ &= hS \sum_{m=0}^\infty \frac{1}{(m+1)!} D^m \nabla H(x)(G(\Phi, x) - x)^m \\ &\quad - hS \sum_{m=1}^\infty \frac{2m}{(m+1)!} D_2^{m-1} Q(x, x)(G(\Phi, x) - x)^m. \end{aligned}$$

Arguing as in the proof of Lemma III.1.9 in [13], we get $hS\overline{\nabla} H(x, \hat{x}) = G(\theta, x)$, with $\theta(\emptyset) = 0$, $\theta(\bullet) = 1$, and

$$\begin{aligned} \theta([\tau_1, \ldots, \tau_m]_\bullet) &= \frac{1}{m+1} \Phi(\tau_1) \cdots \Phi(\tau_m), \\ \theta([\tau_1, \ldots, \tau_m]_\blacktriangle) &= \frac{-2m}{m+1} \Phi(\tau_1) \cdots \Phi(\tau_m). \end{aligned} \tag{5.4}$$

Then we can write (1.9) with $\overline{S}(x, \hat{x}, h)$ given by (5.3) as

$$\begin{aligned} \hat{x} &= x + \sum_{s \in SG} h^n \sum_j b_{sj} \left( \prod_{k=1}^n SR(\phi_{sjk}, x) + (-1)^{|s|_\bullet - 1} \prod_{k=1}^n SR(\phi_{sj(n-k+1)}, x) \right) G(\theta, x) \\ &= x + G(\theta, x) + \sum_{s \in SG,\, n>0} \sum b_j^s \big( G(\phi_{sj1} \diamond \cdots \diamond \phi_{sjn} \diamond \theta, x) + (-1)^{|s|_\bullet - 1} G(\phi_{sjn} \diamond \cdots \diamond \phi_{sj1} \diamond \theta, x) \big) \\ &= G(\Phi, x), \end{aligned}$$

with

$$\Phi = \hat{e} + \theta + \sum_{s\in SG,\ n>0} \sum_j b_{sj} \left(\phi_{sj1} \diamond \cdots \diamond \phi_{sjn} \diamond \theta + (-1)^{|s|_\bullet - 1} \phi_{sjn} \diamond \cdots \diamond \phi_{sj1} \diamond \theta\right). \tag{5.5}$$

□

**Theorem 5.3.** *The discrete gradient method* (1.9) *with* $\overline{S}(x, \hat{x}, h)$ *given by* (5.3) *and* $\overline{\nabla} H \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ *is of order* $p$ *if and only if*

$$\Phi(\tau) = \xi(\tau) \quad \textit{for } |\tau| \le p, \tag{5.6}$$

*where* $\Phi$ *is given by* (5.5) *and the* $\xi$ *is given by* (5.2).

We remark that $\overline{\nabla} H \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ is a necessary condition for the method to be a G-series method for all $S$ and $H$, but not for its order; $\overline{\nabla} H \in C^{p-1}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ is sufficient for the scheme to be of order $p$. The following proposition is presented without its proof, which follows along the lines of the proof of Proposition 3.6.

**Proposition 5.4.** *The* $\Phi$ *of* (5.6) *satisfies*

$$\Phi(\tau) = \hat{e}(\tau) + \theta(\tau) + \sum_{i \text{ s.t. } n\ge 1} \Lambda(\hat{\tau}^i, s^i) \tag{5.7}$$

*where* $\hat{e}(\emptyset) = 1$ *and* $\hat{e}(\tau) = 0$ *for all* $\tau \neq \emptyset$, $\theta$ *is given by* (5.4), *and*

$$\begin{aligned}\Lambda(\hat{\tau}^i, s^i) = \theta([\mu^i_{n+1}]_{s^i_{n+1}}) \big(&\sum_j b_{s^i j} \phi_{s^i j1}(\mu^i_1) \cdots \phi_{s^i jn}(\mu^i_n) \\ &+ (-1)^{|s^i|_\bullet - 1} \sum_j b_{\hat{s}^i j} \phi_{\hat{s}^i jn}(\mu^i_1) \cdots \phi_{\hat{s}^i j1}(\mu^i_n)\big),\end{aligned} \tag{5.8}$$

*with* $\hat{s}^i$ *given by* $\hat{s}^i_k = s^i_{n-k+1}$ *for* $k = 1, \ldots, n$, *and* $\hat{s}^i_{n+1} = s^i_{n+1}$.

As for the AVF method, one does not need to find the order conditions for every tree; it suffices to find the order condition for each energy-preserving linear combination of the form

$$\omega = [\mu_1]_{s_1} \circ [\mu_2]_{s_2} \circ \cdots [\mu_n]_{s_n} \circ [\emptyset]_\bullet + (-1)^n [\mu_n]_{s_n} \circ [\mu_{n-1}]_{s_{n-1}} \circ \cdots [\mu_1]_{s_1} \circ [\emptyset]_\bullet. \tag{5.9}$$

The above does not give every energy-preserving linear combination of the elementary differentials of G-series; it gives the combinations one gets in the scheme (1.9) with $\overline{S}(x, \hat{x}, h)$ given by (5.3). Now, let again $I_l$ and $I_n$ denote the sets of leaf nodes and non-leaf nodes, respectively. If we assume the conditions for order $< p$ to be satisfied, we have an equivalent order condition to (5.6) by

$$\sum_{i\in I_l} \Lambda(\hat{\tau}^i, s^i) = \xi(\tau) - \hat{e}(\tau) - \sum_{i\in I_n} \Lambda(\hat{\tau}^i, s^i), \tag{5.10}$$

where we may use the relation

$$\Lambda(\{\mu^i_1, \ldots, \mu^i_n, \mu^i_{n+1}\}, s^i) = \hat{\theta}([\mu^i_{n+1}]_{s^i_{n+1}}) \Lambda(\{\mu^i_1, \ldots, \mu^i_n, \emptyset\}, \bar{s}^i)$$

to calculate $\Lambda(\hat{\tau}^i)$ for $i \in I_n$. Here $\bar{s}^i$ is $s^i$ with $s^i_{n+1}$ replaced by $\bullet$, and $\hat{\theta}(\emptyset) = 0$, $\hat{\theta}(\bullet) = 1$, and

$$\begin{aligned} \hat{\theta}([\tau_1, \ldots, \tau_m]_\bullet) &= \frac{1}{m+1}\xi(\tau_1)\cdots\xi(\tau_m), \\ \hat{\theta}([\tau_1, \ldots, \tau_m]_\blacktriangle) &= \frac{-2m}{m+1}\xi(\tau_1)\cdots\xi(\tau_m). \end{aligned} \tag{5.11}$$

Note that $\Lambda(\hat{\tau}^1, s^1) = \hat{\theta}(\tau)$.

**Example 5.1.** Consider $\tau = $ , which is part of two combinations of the form (5.9): $\omega = \ + \ $ and $\omega = 2$ . We calculate

$$\begin{aligned} \Lambda(\hat{\tau}^1, s^1) &= \Lambda(\{(\bullet, \ )\}, \blacktriangle) = \hat{\theta}(\ ) = 0, \\ \Lambda(\hat{\tau}^2, s^2) &= \Lambda(\{\ , \emptyset\}, \ ) = \sum_j b_{s^2 j}\phi_{s^2 j1}(\ ) + \sum_j b_{\bar{s}^2 j}\phi_{\bar{s}^2 j1}(\ ) = 2\sum_j b_{s^2 j}\phi_{s^2 j1}(\ ) \\ \Lambda(\hat{\tau}^3, s^3) &= \Lambda(\{\bullet, \bullet\}, \ ) = \hat{\theta}(\ )\Lambda(\{\bullet, \emptyset\}, \ ) = \hat{\theta}(\ )(-\frac{1}{2}\hat{\theta}(\ )) = -1(-\frac{1}{2}(-\frac{4}{3})) = -\frac{2}{3}, \\ \Lambda(\hat{\tau}^4, s^4) &= \Lambda(\{\bullet, \emptyset, \emptyset\}, \ ) = \sum_j b_{s^4 j}\phi_{s^4 j1}(\bullet) + \sum_j b_{\bar{s}^4 j}\phi_{\bar{s}^4 j2}(\bullet) = \sum_{j,k} b_{s^4 j}\phi_{s^4 jk}(\bullet). \end{aligned}$$

Thus (5.10) becomes

$$2\sum_j b_{s^2 j}\phi_{s^2 j1}(\ ) + \sum_{j,k} b_{s^4 j}\phi_{s^4 jk}(\bullet) = \frac{2}{3}$$

for $\tau = $ . We do similar calculations for , and get (5.10) for that to be

$$2\sum_{j,k} b_{s^4 j}\phi_{s^4 jk}(\bullet) = \frac{4}{3}.$$

Thus we have the order condition

$$\sum_{j,k} b_{s^4 j}\phi_{s^4 jk}(\bullet) = \frac{2}{3} \tag{5.12}$$

for $\omega = \ + \ $ , and

$$\sum_j b_{s^2 j}\phi_{s^2 j1}(\ ) = 0$$

for $\omega = 2$ . Note that although the tree gives an energy-preserving elementary differential, this by itself is not of the form (5.9).

From the order conditions in Table 5, we can find an $\overline{S}(x, y, h)$ so that (1.9) becomes a fourth order scheme for any $\overline{\nabla} H \in C^3(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$. For instance, the stem $s = $ has the related order conditions $\sum_j b_{sj} = \frac{1}{2}$ and $\sum_{j,k} b_{sj}\phi_{sjk}(\bullet) = \frac{2}{3}$, which sets the requirements for the term

$$h^2 \sum_j b_{sj}(SQ(x, z_{1j})SQ(x, z_{2j}) + SQ(x, z_{2j})SQ(x, z_{1j}))S.$$

Choosing $b_{s1} = \frac{1}{2}$ and $z_{11} = z_{21} = x + \frac{2}{3}hf(x)$, we have fulfilled these conditions. Likewise, finding terms that satisfy the other order conditions, we get an approximation of $S$ that ensures fourth order convergence, like the $\overline{S}(x, y, h)$ given by (2.17).

| $\lvert\tau\rvert$ | $\omega$ | $s$ | Order condition |
|---|---|---|---|
| 1 | $\bullet$ | $\bullet$ | $\sum_j b_{sj} = \frac{1}{2}$ |
| 2 | | | $\sum_j b_{sj} = \frac{1}{2}$ |
| 3 | | | $\sum_j b_{sj}\phi_{sj1}(\bullet) = \frac{1}{3}$ |
| | | | $\sum_j b_{sj} = \frac{1}{2}$ |
| | | | $\sum_j b_{sj} - \sum_j b_{\bar{s}j} = 0$ |
| | | | $\sum_j b_{sj} = -\frac{1}{24}$ |
| 4 | | | $\sum_j b_{sj}\phi_{sj1}(\bullet)^2 = \frac{1}{4}$ |
| | | | $\sum_j b_{sj}\phi_{sj1}(\;) = 0$ |
| | | | $\sum_j b_{sj}\phi_{sj1}(\;) = \frac{1}{6}$ |
| | | | $\sum_{j,k} b_{sj}\phi_{sjk}(\bullet) = \frac{2}{3}$ |
| | | | $\sum_j b_{sj}\phi_{sj2}(\bullet) - \sum_j b_{\bar{s}j}\phi_{\bar{s}j1}(\bullet) = 0$ |
| | | | $\sum_j b_{sj}\phi_{sj2}(\bullet) - \sum_j b_{\bar{s}j}\phi_{\bar{s}j1}(\bullet) = 0$ |
| | | | $\sum_{j,k} b_{sj}\phi_{sjk}(\bullet) = -\frac{1}{24}$ |
| | | | $\sum_j b_{sj} = \frac{1}{2}$ |
| | | | $\sum_j b_{sj} - \sum_j b_{\bar{s}j} = 0$ |
| | | | $\sum_j b_{sj} - \sum_j b_{\bar{s}j} = 0$ |
| | | | $\sum_j b_{sj} = 0$ |

Table 5: Energy-preserving linear combinations of the form (5.9) and their associated order conditions for the discrete gradient method (1.9) with $\overline{S}(x, \hat{x}, h)$ given by (5.3).

## 5.2 The general case

Allowing for $S$ to be a function of the solution, we define now the set $TV$ of bi-colored trees whose nodes are either circles of triangles, and whose leaves on the cut tree $\tau^b$, defined as the mono-colored tree left when all branches between black and white nodes are cut off, are always circles. Denoting as before black-rooted subtrees by $\tau_i$ and white-rooted subtrees by $\bar{\tau}_i$, the elementary differentials of trees $\tau \in TV$ are given by $F(\bullet)(x) = F(\circ)(x) = f(x) = S\nabla H(x)$ and

$$F(\tau)(x) = \begin{cases} S^{(l)} D^m \nabla H(x)(F(\tau_1)(x), \dots, F(\bar{\tau}_l)(x)) & \text{for } \tau = [\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l,]_{\circledcirc}, \\ S^{(l)} D_2^{m-1} Q(x,x)(F(\tau_1)(x), \dots, F(\bar{\tau}_l)(x)) & \text{for } \tau = [\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l,]_{\triangle}, \end{cases}$$

where $\circledcirc$ can be either $\bullet$ or $\circ$ and $\triangle$ can be either $\blacktriangle$ or $\triangle$. Let $TV_\bullet$ denote the set of trees in $TV$ with black roots, either of the shape $\bullet$ or $\blacktriangle$.

**Definition 5.2.** A V-series is a formal series of the form

$$V(\phi, x) = \phi(\emptyset)x + \sum_{\tau\in TV_\bullet} \frac{h^{|\tau|}}{\sigma(\tau)}\phi(\tau)F(\tau)(x), \tag{5.13}$$

where $\phi : TV_\bullet \cup \{\emptyset\} \to \mathbb{R}$ is an arbitrary mapping, and the symmetry coefficient $\sigma$ is given by (3.1).

Proofs of the theorems in this section can be obtained similarly to the proofs in Chapter 4 and Section 5.1, and are therefore omitted.

We consider now approximations of $S(x)$ that can be written as

$$\begin{aligned}\overline{S}(x,y,h) = \sum_{s\in SG} h^n \sum_j b_{sj}\Bigg(&\prod_{k=1}^{n} S(V(\psi_{sjk},x))R(\phi_{sjk},x)\cdot S(V(\psi_{sj(n+1)},x)) \\ &+(-1)^{|s|_\bullet-1} S(V(\psi_{sj(n+1)},x))\prod_{k=1}^{n} R(\phi_{sj(n-k+1)},x)S(V(\psi_{sj(n-k+1)},x))\Bigg)\end{aligned} \tag{5.14}$$

whenever $y$ is the solution of

$$\frac{y-x}{h} = \overline{S}(x,y,h)\overline{\nabla}H(x,y),$$

with $\phi_{sjk}(\emptyset) = \psi_{sjk}(\emptyset) = 1$ for every $s, j, k$, and with $\sum_j b_{\bullet j} = \frac{1}{2}$.

**Theorem 5.5.** *The discrete gradient scheme* (1.9) *with the approximation of* $S(x)$ *given by* (5.14) *and* $\overline{\nabla}H \in C^\infty(\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^d)$ *is a V-series method. It can be written* $\hat{x} = V(\Phi, x)$, *with*

$$\begin{aligned}\Phi = \hat{e} + \sum_{s\in SG}\sum_j b_{sj}\big(&(\psi_{sj1},\phi_{sj1})\diamond\cdots\diamond(\psi_{sjn},\phi_{sjn})\diamond\hat{\theta}(\psi_{sj(n+1)}) \\ &+(-1)^n(\psi_{sj(n+1)},\phi_{sjn})\diamond\cdots\diamond(\psi_{sj2},\phi_{sj1})\diamond\hat{\theta}(\psi_{sj1})\big).\end{aligned} \tag{5.15}$$

*where*

$$\begin{aligned}\hat{\theta}(a)([\tau_1,\dots,\tau_m,\bar{\tau}_1,\dots,\bar{\tau}_l]_\bullet) &= \frac{1}{m+1}\Phi(\tau_1)\cdots\Phi(\tau_m)a(\bar{\tau}_1)\cdots a(\bar{\tau}_l), \\ \hat{\theta}(a)([\tau_1,\dots,\tau_m,\bar{\tau}_1,\dots,\bar{\tau}_l]_\blacktriangle) &= \frac{-2m}{m+1}\Phi(\tau_1)\cdots\Phi(\tau_m)a(\bar{\tau}_1)\cdots a(\bar{\tau}_l).\end{aligned} \tag{5.16}$$

*The scheme is of order* $p$ *if and only if*

$$\Phi(\tau) = \xi(\tau) \quad \textit{for } |\tau| \le p, \tag{5.17}$$

*where*

$$\xi(\tau) = \begin{cases}\frac{1}{\gamma(\tau)} & \textit{if } \tau\in TP, \\ 0 & \textit{otherwise.}\end{cases} \tag{5.18}$$

As in section 4.2, we cut the branches between black and white nodes, regardless of the shape of the nodes, and denote this tree by $\tau^b$. Number the nodes and reattach the cut-off parts. For the node $i$ and the corresponding stem $s^i$, there exists a unique set of forests $\hat{\tau}^i = \{(\mu_1^i,\eta_1^i),\dots,(\mu_{n+1}^i,\eta_{n+1}^i)\}$ such that

$$\tau = [(\mu_1^i,\eta_1^i)]_{s_1^i}\circ\cdots[(\mu_n^i,\eta_n^i)]_{s_n^i}\circ[(\mu_{n+1}^i,\eta_{n+1}^i)]_{s_{n+1}^i}$$

**Proposition 5.6.** *The* $\Phi$ *of* (5.15) *satisfies*

$$\Phi(\tau) = \hat{e}(\tau) + \sum_{i=1}^{|\tau^b|} \Lambda(\hat{\tau}^i, s^i) \tag{5.19}$$

*where* $\hat{e}(\emptyset) = 1$ *and* $\hat{e}(\tau) = 0$ *for all* $\tau \neq \emptyset$, *and*

$$\begin{aligned}\Lambda(\hat{\tau}^i, s^i) =& \theta([\mu^i_{n+1}]_{s^i_{n+1}})\big(\sum_j b_{s^i j}\psi_{s^i j1}(\eta^i_1)\phi_{s^i j1}(\mu^i_1)\cdots\phi_{s^i jn}(\mu^i_n)\psi_{s^i j(n+1)}(\eta^i_{n+1}) \\ &+ (-1)^{|s^i|_\bullet - 1}\sum_j b_{\hat{s}^i j}\psi_{\hat{s}^i j(n+1)}(\eta^i_1)\phi_{\hat{s}^i jn}(\mu^i_1)\cdots\phi_{\hat{s}^i j1}(\mu^i_n)\psi_{\hat{s}^i j1}(\eta^i_n)\big),\end{aligned} \tag{5.20}$$

*with* $\theta$ *given by* (5.4) *and* $\hat{s}^i$ *given by* $\hat{s}^i_k = s^i_{n-k+1}$ *for* $k = 1, \dots, n$, *and* $\hat{s}^i_{n+1} = s^i_{n+1}$.

The number of trees in $TV$ grows very quickly. However, in our task of finding higher order schemes we may use the lessons of the previous chapters, and require that the arguments of $S$, $\nabla^2 H$ and $Q$ in (5.14) are B-series up to order $p-1$. Then we only need to find order conditions for energy-preserving linear combinations of the form

$$\omega = [(\mu_1, \eta_1)]_{s_1} \circ \cdots \circ [(\mu_n, \eta_n)]_{s_n} \circ [\eta_{n+1}]_\bullet + (-1)^n [(\mu_n, \eta_{n+1})]_{s_n} \circ \cdots \circ [(\mu_1, \eta_2)]_{s_1} \circ [\eta_1]_\bullet, \tag{5.21}$$

where $\mu_i$ and $\eta_i$ are forests of trees in $TP_\bullet$ and $TP_\circ$ respectively, for $i = 1, \dots, n+1$. Thus we can disregard any tree with $\vartriangle$ in it. Furthermore, we may color all nodes of the trees in $\mu_i$ and $\eta_i$ except the roots gray, and let the elementary differentials corresponding to these trees be the same as the elementary differentials of B-trees.

We find the order conditions

$$\sum_{i \in I_l} \Lambda(\hat{\tau}^i, s^i) = \xi(\tau) - \hat{e}(\tau) - \sum_{i \in I_n} \Lambda(\hat{\tau}^i, s^i),$$

by using the relation

$$\Lambda(\{(\mu^i_1, \eta^i_1), \dots, (\mu^i_{n+1}, \eta^i_{n+1})\}, s^i) = \hat{\theta}([\mu^i_{n+1}]_{s^i_{n+1}})\Lambda(\{(\mu^i_1, \eta^i_1), \dots, (\emptyset, \eta^i_{n+1})\}, \bar{s}^i)$$

to calculate $\Lambda(\hat{\tau}^i)$ for $i \in I_n$. The $\hat{\theta}$ is given by (5.11), and $\bar{s}^i$ is $s^i$ with $s^i_{n+1}$ replaced by $\bullet$.

**Example 5.2.** Consider , which is part of the energy-preserving linear combination $+$ . We have two black nodes, and calculate

$$\begin{aligned}\Lambda(\hat{\tau}^1, s^1) &= \Lambda(\{(\bullet, \;)\}, \blacktriangle) = \hat{\theta}(\;)\Lambda(\{(\emptyset, \;)\}, \bullet) = -\xi(\bullet)\xi(\;) = -\frac{1}{6}, \\ \Lambda(\hat{\tau}^2, s^2) &= \Lambda(\{(\emptyset, \;), (\emptyset, \emptyset)\}, \;) = \sum_j b_{s^2 j}\psi_{s^2 j1}(\;) + \sum_j b_{\bar{s}^2 j}\psi_{\bar{s}^2 j2}(\;) = \sum_j b_{s^2 j}(\psi_{s^2 j1} + \psi_{s^2 j2})(\;).\end{aligned}$$

Hence the order condition associated to this linear combination is

$$\sum_j b_{s^2 j}(\psi_{s^2 j1} + \psi_{s^2 j2})(\;) = \frac{1}{6}.$$

| $\lvert\tau\rvert$ | $\omega$ | $s$ | Order condition |
|---|---|---|---|
| 1 | $\bullet$ | $\bullet$ | $\sum_j b_{sj} = \frac{1}{2}$ |
| 2 | | $\bullet$ | $2\sum_j b_{sj}\psi_{sj1}(\bullet) = \frac{1}{2}$ |
| | | | $\sum_j b_{sj} = \frac{1}{2}$ |
| 3 | | $\bullet$ | $2\sum_j b_{sj}\psi_{sj1}(\bullet)^2 = \frac{1}{3}$ |
| | | $\bullet$ | $2\sum_j b_{sj}\psi_{sj1}(\,) = \frac{1}{6}$ |
| | | | $\sum_j b_{sj}\phi_{sj1}(\bullet) = \frac{1}{3}$ |
| | | | $\sum_j b_{sj}(\psi_{sj1}+\psi_{sj2})(\bullet) = \frac{1}{2}$ |
| | | | $\sum_j b_{sj}(\psi_{sj2}-\psi_{sj1})(\bullet) = -\frac{1}{12}$ |
| | | | $\sum_j b_{sj} = \frac{1}{2}$ |
| | | | $\sum_j b_{sj} - \sum_j b_{\bar{s}j} = 0$ |
| | | | $\sum_j b_{sj} = -\frac{1}{24}$ |
| 4 | | $\bullet$ | $2\sum_j b_{sj}\psi_{sj1}(\bullet)^3 = \frac{1}{4}$ |
| | | $\bullet$ | $2\sum_j b_{sj}\psi_{sj1}(\bullet)\psi_{0j1}(\,) = \frac{1}{8}$ |
| | | $\bullet$ | $2\sum_j b_{sj}\psi_{sj1}(\,) = \frac{1}{12}$ |
| | | $\bullet$ | $2\sum_j b_{sj}\psi_{sj1}(\,) = \frac{1}{24}$ |
| | | | $\sum_j b_{sj}\phi_{sj1}(\bullet)^2 = \frac{1}{4}$ |
| | | | $\sum_j b_{sj}\phi_{sj1}(\,) = \frac{1}{6}$ |
| | | | $\sum_j b_{sj}\psi_{sj1}(\bullet)\psi_{sj2}(\bullet) = \frac{1}{8}$ |
| | | | $\sum_j b_{sj}\phi_{sj1}(\bullet)(\psi_{sj1}+\psi_{sj2})(\bullet) = \frac{1}{3}$ |
| | | | $\sum_j b_{sj}(\psi_{sj1}(\bullet)^2+\psi_{sj2}(\bullet)^2) = \frac{1}{3}$ |
| | | | $\sum_j b_{sj}(\psi_{sj1}+\psi_{sj2})(\,) = \frac{1}{6}$ |
| | | | $\sum_j b_{sj}\phi_{sj1}(\bullet)(\psi_{sj2}-\psi_{sj1})(\bullet) = -\frac{1}{24}$ |
| | | | $\sum_j b_{sj}(\psi_{sj2}(\bullet)^2-\psi_{sj1}(\bullet)^2) = -\frac{1}{12}$ |
| | | | $\sum_j b_{sj}(\psi_{sj2}-\psi_{sj1})(\,) = -\frac{1}{24}$ |
| | | | $\sum_{j,k} b_{sj}\phi_{sjk}(\bullet) = \frac{2}{3}$ |
| | | | $\sum_j b_{sj}\psi_{sj2}(\bullet) = \frac{2}{3}$ |
| | | | $\sum_j b_{sj}(\psi_{sj1}+\psi_{sj3})(\bullet) = \frac{1}{2}$ |
| | | | $\sum_j b_{sj}\phi_{sj2}(\bullet) - \sum_j b_{\bar{s}j}\phi_{\bar{s}j1}(\bullet) = 0$ |

Table 6: Energy-preserving linear combinations of the form (5.9). Continued in Table 7.

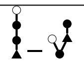
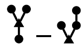
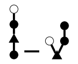
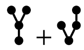
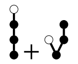
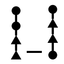
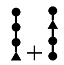
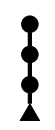
| $\lvert\tau\rvert$ | $\omega$ | $s$ | Order condition |
|---|---|---|---|
| 4 |  |  | $\sum_j b_{sj}(\psi_{sj1}-\psi_{sj3})(\bullet)=\frac{1}{12}$ |
| |  |  | $\sum_j b_{sj}\phi_{sj2}(\bullet)-\sum_j b_{\bar{s}j}\phi_{\bar{s}j1}(\bullet)=0$ |
| |  |  | $\sum_j b_{sj}\psi_{sj3}(\bullet)-\sum_j b_{\bar{s}j}\psi_{\bar{s}j1}(\bullet)=0$ |
| |  |  | $\sum_{j,k} b_{sj}\phi_{sjk}(\bullet)=-\frac{1}{24}$ |
| |  |  | $2\sum_j b_{2j}\psi_{2j1}(\bullet)=-\frac{1}{24}$ |
| |  |  | $\sum_j b_{sj}(\psi_{sj1}+\psi_{sj3})(\bullet)=-\frac{1}{24}$ |
| |  |  | $\sum_j b_{sj}=\frac{1}{2}$ |
| |  |  | $\sum_j b_{sj}-\sum_j b_{\bar{s}j}=0$ |
| |  |  | $\sum_j b_{sj}=-\frac{1}{12}$ |
| |  |  | $\sum_j b_{sj}=0$ |

Table 7: Energy-preserving linear combinations of the form (5.9). Continuing from Table 6.

We consider the order conditions for trees with $|\tau|\le 3$ displayed in Table 6, and find that

$$\begin{aligned}\overline{S}(x,\cdot,h)=&\frac{1}{4}S(x)+\frac{3}{4}S(z_3)\\&+hS(z_2)Q(x,z_3)S(z_2)+\frac{1}{4}h\big(S(z_1)\nabla^2H(x)S(x)-S(x)\nabla^2H(x)S(z_1)\big)\\&+h^2S(x)Q(x,x)S(x)Q(x,x)S(x)-\frac{1}{12}h^2\,S(x)\nabla^2H(x)S(x)\nabla^2H(x)S(x),\end{aligned}\tag{5.22}$$

where

$$z_1=x+\frac{1}{3}hf(x),\qquad z_2=x+\frac{1}{2}hf(x),\qquad z_3=x+\frac{2}{3}hf(z_1),$$

guarantees third order convergence of the scheme (1.9) if $\overline{\nabla}H(x)\in C^2(\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^d)$. An approximation of $S(x)$ satisfying all the order conditions in tables 6 and 7 is given by

$$\begin{aligned}
\overline{S}(x,\cdot,h) =& \frac{1}{2}(S(z_{11}+z_{12})+S(z_{11}-z_{12})) + \frac{1}{12}h\big(S(z_6)\nabla^2 H(z_2)S(x) - S(x)\nabla^2 H(z_2)S(z_6)\big)\\
&+ \frac{3}{7}h\big(S(z_3)Q(x,z_5)S(z_4)+S(z_4)Q(x,z_5)S(x,z_3)\big)\\
&+ \frac{8}{105}hS(x)Q(x,z_7)S(x) + \frac{1}{15}hS(x)Q(x,x)S(x)\\
&+ h^2 S(z_2)Q(x,z_5)S(z_8)Q(x,z_5)S(z_2)\\
&- \frac{1}{12}h^2 S(z_2)\nabla^2 H(z_2)S(z_2)\nabla^2 H(z_2)S(z_2)\\
&+ \frac{1}{6}h^2(S(z_2)-S(x))\nabla^2 H(x)S(x)Q(x,x)S(x)\\
&- \frac{1}{6}h^2 S(x)Q(x,x)S(x)\nabla^2 H(x)(S(z_2)-S(x))\\
&+ h^3 S(x)Q(x,x)S(x)Q(x,x)S(x)Q(x,x)\\
&- \frac{1}{12}h^3 S(x)\nabla^2 H(x)S(x)\nabla^2 H(x)S(x)Q(x,x)S(x)\\
&- \frac{1}{12}h^3 S(x)Q(x,x)S(x)\nabla^2 H(x)S(x)\nabla^2 H(x)S(x),
\end{aligned} \tag{5.23}$$

with

$$\begin{aligned}
z_1 &= x+\frac{1}{3}hf(x), & z_5 &= x+\frac{2}{3}hf(z_2), & z_9 &= x+hf(z_6),\\
z_2 &= x+\frac{1}{2}hf(x), & z_6 &= x+hf(z_2), & z_{10} &= x+hf(z_9),\\
z_3 &= x+\frac{7-\sqrt{7}}{12}hf(z_1), & z_7 &= x+\frac{5}{4}hf(z_2), & z_{11} &= \frac{1}{3}(x+z_2+z_6)+\frac{1}{12}(-z_9+z_{10}),\\
z_4 &= x+\frac{7+\sqrt{7}}{12}hf(z_1), & z_8 &= x+\frac{4}{3}hf(z_2), & z_{12} &= \frac{\sqrt{3}}{36}(7x-2z_2-4z_6+z_9-2z_{10}),
\end{aligned}$$

and hence a discrete gradient scheme with this $\overline{S}(x,\cdot,h)$ and any $\overline{\nabla}H(x,y) \in C^3(\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^d)$ will be of fourth order.

One advantage of choosing the AVF discrete gradient is that the resulting scheme generally requires fewer computations at each time step. This is clearly evident in the above example: if $\overline{\nabla} = \overline{\nabla}_{\text{AVF}}$, then (5.22) collapses to (4.17), and (5.23) collapses to (4.19). However, if the AVF discrete gradient is difficult to calculate, there can also be much to gain in computational cost by choosing a symmetric discrete gradient, like the symmetrized Itoh–Abe discrete gradient (2.12) or the Furihata discrete gradient (2.14). Then one can ignore the order condition for any combination (5.21) for which $s_j = \blacktriangle$ and $\mu_j = \emptyset$ for some $j \in [1,n]$, since this corresponds to elementary differentials involving $Q(x,x)$, which we recall is zero when the discrete gradient is symmetric. If we consider the conditions for fourth order presented in tables 6 and 7, this eliminates 17 of the 22 conditions for trees with $\blacktriangle$ in the stem. By considering the remaining order conditions we get that, if $\overline{\nabla}H(x,y) \in C^3(\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^d)$ and $\overline{\nabla}H(x,y) = \overline{\nabla}H(y,x)$, the discrete gradient scheme (1.9) is of fourth order if

$$\begin{aligned}\overline{S}(x,\cdot,h) =& \frac{1}{2}(S(z_5+z_6)+S(z_5-z_6)) + \frac{1}{12}h\left(S(z_2)\nabla^2 H(z_1)S(x) - S(x)\nabla^2 H(z_1)S(z_2)\right)\\ &+\frac{8}{9}hS(z_1)Q(z_7)S(z_1) - \frac{1}{12}h^2 S(z_1)\nabla^2 H(z_1)S(z_1)\nabla^2 H(z_1)S(z_1),\end{aligned}\tag{5.24}$$

with

$$\begin{aligned}
&z_1 = x+\frac{1}{2}hf(x), && z_3 = x+hf(z_2), && z_5 = \frac{1}{3}(x+z_1+z_2)+\frac{1}{12}(-z_3+z_4),\\
&z_2 = x+hf(z_1), && z_4 = x+hf(z_3), && z_6 = \frac{\sqrt{3}}{36}(7x-2z_1-4z_2+z_3-2z_4),\\
&z_7 = x+\frac{3}{4}hf(z_1).
\end{aligned}$$

If $S$ is constant, (5.24) simplifies to

$$\overline{S}(x,\cdot,h) = S + \frac{8}{9}hSQ(z_7)S - \frac{1}{12}h^2 S\nabla^2 H(z_1)S\nabla^2 H(z_1)S. \tag{5.25}$$

# 6 Numerical experiments and conclusions

The Hénon–Heiles system can be written on the form (1.6) with

$$S = \begin{pmatrix} 0 & I\\ -I & 0\end{pmatrix}, \qquad H(q,p) = \frac{1}{2}(q_1^2+q_2^2+p_1^2+p_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3, \tag{6.1}$$

where $I$ is the $2\times 2$ identity matrix. We use here the same initial conditions used in [26]: $q_1 = \frac{1}{10}$, $q_2 = -\frac{1}{2}$, $p_1 = p_2 = 0$. The order of some of the energy-preserving methods proposed in this paper are confirmed by the left plot in Figure 1. We compare the performance of the fourth order discrete gradient methods obtained by using the $\overline{S}$ given by (2.17) coupled with three different discrete gradients: the Itoh–Abe discrete gradient (2.11), the Furihata discrete gradient (2.14), and the AVF discrete gradient (1.10). The symmetrized Itoh–Abe discrete gradient (2.12) is for this $H$ identical to the Furihata discrete gradient. The AVF and Furihata discrete gradient methods perform in this case very similarly, and thus the error from the Furihata discrete gradient method is excluded from the right plot in Figure 1. We observe that, although it initially performs on par with the AVF method, the Itoh–Abe discrete gradient method gives a lower global error than the other fourth order methods as time goes on. Note however that this method requires the most computations at every time step.

The methods should also be tested on a skew-gradient system with non-constant $S$. We choose the Lotka–Volterra system also used for numerical experiments in [5]. It is given by

$$S = \frac{1}{2}\begin{pmatrix} 0 & -x_1x_2 & x_1x_3\\ x_1x_2 & 0 & -2x_2x_3\\ -x_1x_3 & 2x_2x_3 & 0\end{pmatrix}, \qquad H(x) = 2x_1+x_2+2x_3+\ln(x_2)-2\ln(x_3), \tag{6.2}$$

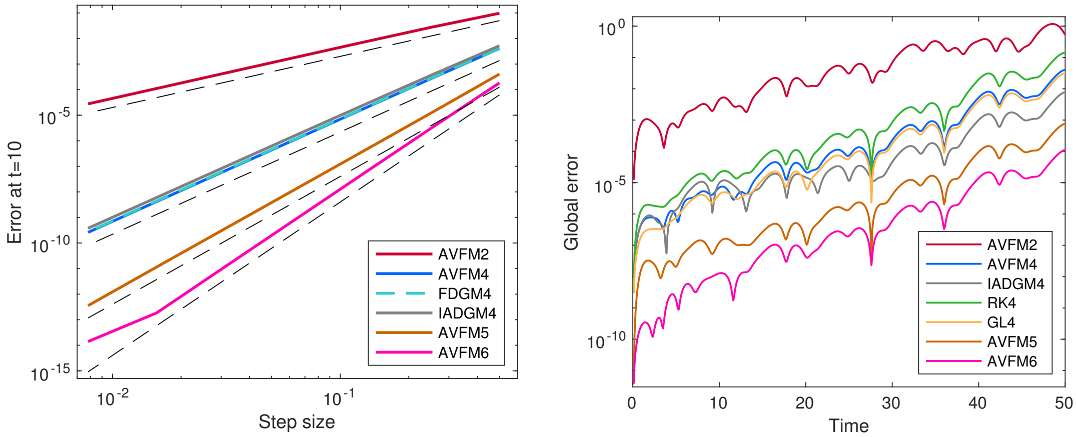



Figure 1: Error plots for the Hénon–Heiles system (6.1) solved by various discrete gradient methods: AVFM2 is the standard AVF method (1.11); AVFM4 is the AVF discrete gradient method with $\overline{S}$ given by (2.18); FDGM4 is the Furihata discrete gradient method with $\overline{S}$ given by (5.25); IADGM4 is the Itoh–Abe discrete gradient method with $\overline{S}$ given by (2.17); AVFM5 is the scheme (3.19); AVFM6 is (3.21). RK4 is the classic Runge–Kutta method and GL4 is the fourth order Gauss–Legendre method, included for comparison. The black dashed lines in the order plot are reference lines of order two, four, five and six. The step size in the right plot is $h = 0.1$.

and initial conditions $x_1 = 1$, $x_2 = \frac{19}{10}$, $x_3 = \frac{1}{2}$. For this $H$, the Itoh–Abe, Furihata and AVF discrete gradients are all equivalent. We consider fourth order discrete gradient methods where $\nabla S$ is given either dependent on or independent of $\hat{x}$; that is, (4.18) or (4.19). The implicitly given (4.19) yields a significantly lower error in the solution of the corresponding discrete gradient method, as can be witnessed from the left plot in Figure 3. In contrast to what we observed for the canonical Hamiltonian system studied above, none of the discrete gradient methods give a global error lower than that of the fourth order Gauss–Legendre method.

The main purpose of this paper has been to develop order theory for discrete gradient methods, rather than the development of specific schemes. Hence we have simply proposed some higher order schemes satisfying the derived order conditions; analysis to find more optimal schemes is something we leave for the future. After such an analysis is performed, the methods could be tested on more advanced problems than those considered above, e.g. for the temporal discretization of Hamiltonian partial differential equations, and their performance as measured by accuracy relative to computational cost could be compared to existing methods.

The order theory presented here can possibly be developed further in a couple of different directions. The schemes given in this paper with $\overline{S}$ independent of $\hat{x}$ are linearly implicit when $H$ is quadratic; if the order theory is extended to the polarized discrete gradient methods of [18, 6], we could get higher order linearly implicit multi-step schemes for systems with polynomial first integrals of any degree. Another avenue could be to consider order conditions for the discrete Riemannian gradient methods presented in [1]. Then the results in the previous chapter are especially interesting, since the integral in the AVF discrete Riemannian gradient can be challenging to compute analytically.

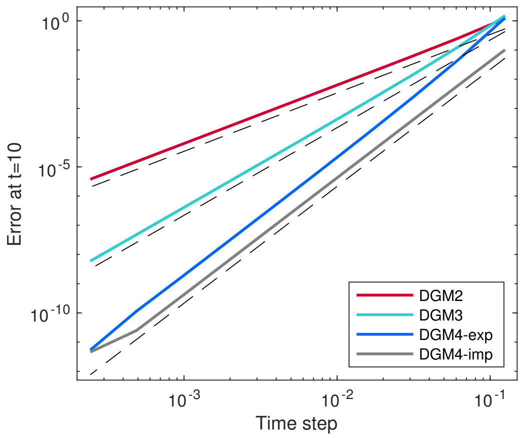



Figure 2: Order or discrete gradient methods applied to the Lotka–Volterra system (6.2), with different $\overline{S}$: $\overline{S}(x,\hat{x}) = S(\frac{x+\hat{x}}{2})$ for DGM2, (4.17) for DGM3, (4.19) for DGM4-exp, (4.18) for DGM4-imp. The dashed lines are reference lines of order two, three and four.

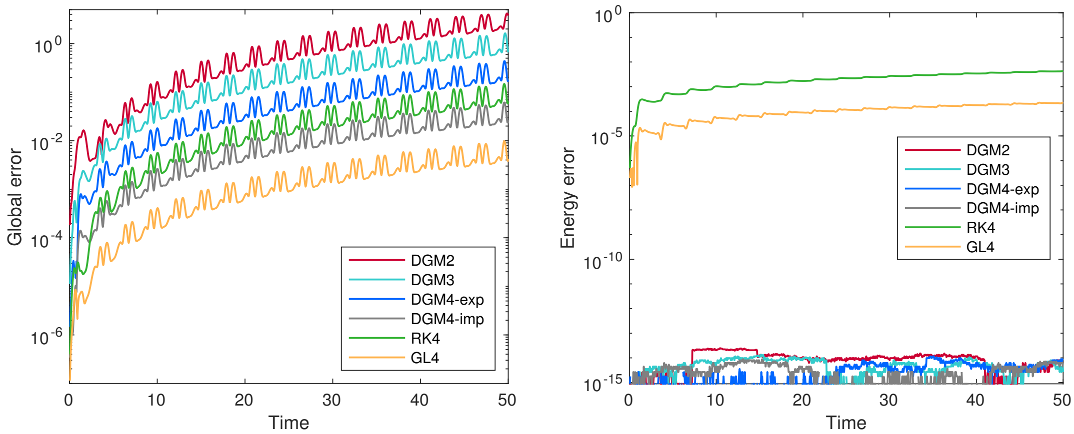



Figure 3: Error in the solution and in the energy for discrete gradient methods with $\overline{S}$ given by $\overline{S}(x,\hat{x}) = S(\frac{x+\hat{x}}{2})$ for DGM2, (4.17) for DGM3, (4.19) for DGM4-exp and (4.18) for DGM4-imp, applied to the Lotka–Volterra system, with step size $h = 0.05$. For comparison, errors from using the standard fourth order Runge–Kutta (RK4) and Gauss–Legendre (GL4) methods are also included.

**Acknowledgments**

The author is grateful to Brynjulf Owren for the invaluable suggestions he has provided during the writing of this paper. This research was supported by the Research Council of Norway, project no. 231632. Some of the work was undertaken while the author was a guest at the Isaac Newton Institute for Mathematical Sciences in Cambridge.

## References

[1] E. Celledoni, S. Eidnes, B. Owren, and T. Ringholm. Energy-preserving methods on Riemannian manifolds. *Math. Comp.*, 89(322):699–716, 2020.

[2] E. Celledoni, R. I. McLachlan, D. I. McLaren, B. Owren, G. R. W. Quispel, and W. M. Wright. Energy-preserving Runge-Kutta methods. *M2AN Math. Model. Numer. Anal.*, 43(4):645–649, 2009.

[3] P. Chartier, E. Faou, and A. Murua. An algebraic approach to invariant preserving integrators: the case of quadratic and Hamiltonian invariants. *Numer. Math.*, 103(4):575–590, 2006.

[4] A. J. Chorin, M. F. McCracken, T. J. R. Hughes, and J. E. Marsden. Product formulas and numerical algorithms. *Comm. Pure Appl. Math.*, 31(2):205–256, 1978.

[5] D. Cohen and E. Hairer. Linear energy-preserving integrators for Poisson systems. *BIT*, 51(1):91–101, 2011.

[6] M. Dahlby and B. Owren. A general framework for deriving integral preserving numerical methods for PDEs. *SIAM J. Sci. Comput.*, 33(5):2318–2340, 2011.

[7] S. Eidnes, B. Owren, and T. Ringholm. Adaptive energy preserving methods for partial differential equations. *Adv. Comput. Math.*, 44(3):815–839, 2018.

[8] E. Faou, E. Hairer, and T.-L. Pham. Energy conservation with non-symplectic methods: examples and counter-examples. *BIT*, 44(4):699–709, 2004.

[9] D. Furihata. Finite difference schemes for $\partial u/\partial t = (\partial/\partial x)^{\alpha}\delta G/\delta u$ that inherit energy conservation or dissipation property. *J. Comput. Phys.*, 156(1):181–205, 1999.

[10] D. Furihata and T. Matsuo. *Discrete variational derivative method*. Chapman & Hall/CRC Numerical Analysis and Scientific Computing. CRC Press, Boca Raton, FL, 2011. A structure-preserving numerical method for partial differential equations.

[11] O. Gonzalez. Time integration and discrete Hamiltonian systems. *J. Nonlinear Sci.*, 6(5):449–467, 1996.

[12] E. Hairer. Energy-preserving variant of collocation methods. *JNAIAM. J. Numer. Anal. Ind. Appl. Math.*, 5(1-2):73–84, 2010.

[13] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[14] A. Harten, P. D. Lax, and B. van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.*, 25(1):35–61, 1983.

[15] T. Itoh and K. Abe. Hamiltonian-conserving discrete canonical equations based on variational difference quotients. *J. Comput. Phys.*, 76(1):85–102, 1988.

[16] Y. A. Kriksin. A conservative difference scheme for a system of Hamiltonian equations with external action. *Zh. Vychisl. Mat. i Mat. Fiz.*, 33(2):206–218, 1993.

[17] J. E. Marsden. *Lectures on mechanics*, volume 174 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1992.

[18] T. Matsuo and D. Furihata. Dissipative or conservative finite-difference schemes for complex-valued nonlinear partial differential equations. *J. Comput. Phys.*, 171(2):425–447, 2001.

[19] T. Matsuo, M. Sugihara, D. Furihata, and M. Mori. Spatially accurate dissipative or conservative finite difference schemes derived by the discrete variational method. *Japan J. Indust. Appl. Math.*, 19(3):311–330, 2002.

[20] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux. Geometric integration using discrete gradients. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357(1754):1021–1045, 1999.

[21] D. I. McLaren and G. R. W. Quispel. Integral-preserving integrators. *J. Phys. A*, 37(39):L489–L495, 2004.

[22] D. I. McLaren and G. R. W. Quispel. Bootstrapping discrete-gradient integral-preserving integrators to fourth order. In M. Daniel and S. Rajasekar, editors, *Nonlinear dynamics*, pages 157–171. Narosa Publishing House, 2009.

[23] S. P. Nørsett and A. Wolfbrandt. Order conditions for Rosenbrock type methods. *Numer. Math.*, 32(1):1–15, 1979.

[24] R. A. Norton, D. I. McLaren, G. R. W. Quispel, A. Stern, and A. Zanna. Projection methods and discrete gradient methods for preserving first integrals of ODEs. *Discrete Contin. Dyn. Syst.*, 35(5):2079–2098, 2015.

[25] R. A. Norton and G. R. W. Quispel. Discrete gradient methods for preserving a first integral of an ordinary differential equation. *Discrete Contin. Dyn. Syst.*, 34(3):1147–1170, 2014.

[26] G. R. W. Quispel and D. I. McLaren. A new class of energy-preserving numerical integration methods. *J. Phys. A*, 41(4):045206, 7, 2008.

[27] T. Yaguchi, T. Matsuo, and M. Sugihara. The discrete variational derivative method based on discrete differential forms. *J. Comput. Phys.*, 231(10):3963–3986, 2012.