PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

# Order conditions for numerical integrators obtained by composing simpler integrators

A. Murua and J.M. Sanz–Serna

*Phil. Trans. R. Soc. Lond. A* 1999 **357**, doi: 10.1098/rsta.1999.0365, published 15 April 1999

THE ROYAL SOCIETY

# Order conditions for numerical integrators obtained by composing simpler integrators

By A. Murua[1] and J. M. Sanz-Serna[2]

[1]*Konputazio Zientziak eta A. A., Informatika Fakultatea, EHU/UPV, Donostia/San Sebastián, Spain*

[2]*Departamento de Matemática Aplicada y Computación, Facultad de Ciencias, Universidad de Valladolid, Valladolid, Spain*

For numerical one-step integration methods obtained by composing or concatenating simpler methods, we study the conditions that the method has to satisfy to attain a prescribed order of accuracy. An existing methodology uses the Baker–Campbell–Hausdorff formula; we develop an alternative technique based on the use of rooted trees and similar to that which is standard in the analysis of Runge–Kutta methods. In the present approach, the order conditions can be written down easily by transcribing the structure of the corresponding rooted trees.

**Keywords: geometric integration; splitting methods; order conditions; composition methods; rooted trees; one-step methods**

## 1. Introduction

The purpose of this paper is to show how to write easily the order conditions for one-step integrators obtained by composing simpler methods. Such composition methods have gained prominence in recent years, mainly through their use in geometric integration (Sanz-Serna 1997).

Let

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(y) \tag{1.1}$$

be the system to be integrated, with the function $y$ taking values in $\mathbb{R}^D$. Each one-step method for the integration of (1.1) is described by a mapping $\psi_{h,f} : \mathbb{R}^D \to \mathbb{R}^D$, so that, if $y_n$ is the numerical solution at time $t_n$, then $y_{n+1} = \psi_{h,f}(y_n)$ is the numerical solution at the next time level $t_{n+1} = t_n + h$. For instance, $\psi_{h,f}(y) = y + hf(y)$ corresponds to Euler's rule. A method is of *order* $r$ if the local error $\psi_{h,f}(y) - \phi_{h,f}(y)$ is $O(h^{r+1})$ as $h \to 0$, where $\phi_{t,f}$ denotes the exact solution flow at time $t$ of the system (1.1). *Consistency* means order greater than or equal to 1. When there is no ambiguity as to the system being integrated, we write $\psi_h$ rather than $\psi_{h,f}$ and $\phi_h$ rather than $\phi_{h,f}$.

The most classical family of composition integrators operates in the case where $f$ in (1.1) can be written as $f = f_1 + f_2$ and each of the split systems

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f_1(y), \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = f_2(y)$$

can be integrated in closed form. Then the composition or concatenation

$$\psi_{h,f} = \phi_{h,f_2} \circ \phi_{h,f_1} \tag{1.2}$$

defines a consistent *split-step* method for the integration of (1.1). The celebrated second-order splitting method due to Strang (1968) is given by

$$\psi_{h,f} = \phi_{h/2,f_2} \circ \phi_{h,f_1} \circ \phi_{h/2,f_2}, \tag{1.3}$$

and, in general, one may consider methods of the form

$$\phi_{h,f} = \phi_{b_s h,f_2} \circ \phi_{a_s h,f_1} \circ \cdots \circ \phi_{b_2 h,f_2} \circ \phi_{a_2 h,f_1} \circ \phi_{b_1 h,f_2} \circ \phi_{a_1 h,f_1} \circ \phi_{b_0 h,f_2}, \tag{1.4}$$

where the method coefficients $a_i$, $b_i$ may be chosen to boost the order. In (1.4) we have assumed for later convenience that the first and last substeps involve $f_2$; this is no restriction because we can take $b_0 = 0$ and or $b_s = 0$.

A second type of composition integrators consists of methods of the form

$$\psi_h = \psi^{[\mathrm{B}]}_{\gamma_s h} \circ \cdots \circ \psi^{[\mathrm{B}]}_{\gamma_2 h} \circ \psi^{[\mathrm{B}]}_{\gamma_1 h}, \tag{1.5}$$

where the $\gamma_i$ are real constants and $\psi^{[\mathrm{B}]}_h$, the so-called basic method, is a given consistent method for the integration of (1.1). In practice $\psi^{[\mathrm{B}]}_h$ is a low-order method (typically $r = 2$) with a favourable geometric property (for instance $\psi^{[\mathrm{B}]}_h$ is symplectic for Hamiltonian systems (Sanz-Serna & Calvo 1994), or volume preserving for divergence-free systems); one aims at choosing $s$ and the $\gamma_i$s so as to obtain a high-order method (1.5), which will automatically inherit the relevant geometric property if this property is preserved by composition. Methods of the form (1.5) have been considered frequently in the recent literature; a very well-known reference is Yoshida (1990). One usually works under the assumption that the basic method is *self-adjoint*. Recall that the *adjoint* of a method $\psi_h$ is, by definition, the method $\psi^*_h$ such that $\psi^*_{-h} \circ \psi_h = \mathrm{id}$ (id denotes the identity mapping); a method is self-adjoint if $\psi^*_h$ coincides with $\psi_h$, i.e. if $\psi_{-h} \circ \psi_h = \mathrm{id}$. When the basic method is self-adjoint it is possible to choose $s$ and the $\gamma_i$ for (1.5) to have any desired order $r$ (see, among others, Yoshida (1990) and § 13.1 of Sanz-Serna & Calvo (1994)).

When the basic method is not self-adjoint the format (1.5) turns out not to be too advantageous and it is better to consider compositions of the more general form (McLachlan 1995)

$$\psi_h = \psi^{[\mathrm{B}]}_{\beta_s h} \circ \psi^{[\mathrm{B}]*}_{\alpha_s h} \circ \cdots \circ \psi^{[\mathrm{B}]}_{\beta_2 h} \circ \psi^{[\mathrm{B}]*}_{\alpha_2 h} \circ \psi^{[\mathrm{B}]}_{\beta_1 h} \circ \psi^{[\mathrm{B}]*}_{\alpha_1 h}. \tag{1.6}$$

A list of references for composition methods can be found in Sanz-Serna (1997). For references in the chemistry literature, see Sans-Serna & Portillo (1996).

In any of the formats (1.4), (1.5) or (1.6), the need arises to find the *order conditions,* i.e. the equations in the parameters $a_i$, $b_i$, $\gamma_i$ or $\alpha_i$, $\beta_i$ that ensure that the composition method has a targeted order $r$. The standard methodology for deriving such order conditions for composition methods involves the use of the Baker–Campbell–Hausdorff (BCH) formula (Bourbaki 1989; Hairer *et al.* 1993) which has been reviewed in detail in Sanz-Serna (1997).

In the case of (1.6) this methodology involves the following steps (the cases (1.4), (1.5) are similar and will not be discussed).

(i) Write the *modified* vector field $\tilde{f}^{[\mathrm{B}]}_h$ of the basic method. This is a formal series

$$\tilde{f}^{[\mathrm{B}]}_h(y) = F^{(1)}(y) + hF^{(2)}(y) + h^2 F^{(3)}(y) + \cdots \tag{1.7}$$

such that (at least as a formal series) $\psi^{[\mathrm{B}]}_h$ coincides with the $h$-flow $\phi_{h,\tilde{f}^{[\mathrm{B}]}_h}$. Note that, from consistency, $F^{(1)} = f$ and that $\psi^{[\mathrm{B}]}_h$ has order $r \geqslant 2$ iff $F^{(2)} = \cdots = F^{(r)} = 0$.

The modified vector field $\tilde{f}_h^{[\mathrm{B}]*}$ of the adjoint method is obtained by changing $h$ into $-h$, i.e.

$$\tilde{f}_h^{[\mathrm{B}]*}(y) = F^{(1)}(y) - hF^{(2)}(y) + h^2F^{(3)}(y) - \cdots . \tag{1.8}$$

(ii) After step (i), each of the mappings in the right-hand side of (1.6) may be rewritten as a flow. In turn, each flow can be identified with an *exponential* of a Lie operator.

(iii) All the exponentials in the right-hand side of (1.6) are combined into a single exponential of a Lie operator by repeatedly using the BCH formula. This is a formula $C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] + \cdots$ for writing the product $\exp(A)\exp(B) = (I + A + \frac{1}{2}A^2 + \cdots)(I + B + \frac{1}{2}B^2 + \cdots)$ as an exponential $\exp(C)$ (square brackets denote Lie brackets or commutators).

(iv) The exponential found in (iii) is interpreted as a flow of a vector field $\tilde{f}_h$. Then, the overall method $\psi_h$ coincides with $\phi_{h,\tilde{f}_h}$, and therefore $\tilde{f}_h$ is in fact the modified vector field of $\psi_h$. Once this modified vector field is available, we impose the condition that it differs from $f$ in terms $O(h^r)$ to ensure that the composition method has order $r$.

When following this methodology, the expression for $\tilde{f}_h$ that one finds is a power series $\sum C_i h^i$ where the $C_i$ are complicated elements of the Lie algebra generated by the functions $F^{(i)}$ in (1.7), i.e. the $C_i$ are obtained from the functions $F^{(i)}$ by repeatedly computing linear combinations and Lie brackets. Thus, to annihilate a given $C_i$ in step (iv), we must first identify a basis of the Lie algebra (which is an easy standard task if this algebra is free (Bourbaki 1989)) and then decompose $C_i$ in terms of the chosen basis (which is a complicated job). The order conditions are obtained by equating to zero the coefficients of the $C_i$ in the chosen basis.

Due to the complexity inherent in the BCH formula itself and to the difficulties in writing the expressions for the $C_i$ in terms of the chosen basis, the methodology above is, except for very low orders $r$, difficult to carry out by hand computation. This is similar to the situation for Runge–Kutta (RK) methods before the sixties, when deriving by hand the order conditions for, say, $r = 4$ was a major task. In the RK case the introduction (mainly by Butcher (1963, 1964)) of a simple formalism based on the use of rooted trees has made it possible to easily write the order conditions for arbitrarily high $r$ (Butcher 1987; Hairer *et al.* 1993; Sanz-Serna & Calvo 1994). Our aim in this paper is to introduce a similar formalism for composition methods.

We now briefly describe the contents of the present article. Most of the paper deals with methods of the form (1.6); the cases of compositions (1.5) and (1.4) are treated separately in the final §§ 6 and 7. For methods (1.6), we begin by finding the expansions in powers of $h$ of $\psi_h$ and $\phi_h$ in terms of rooted trees and elementary differentials in an RK-like approach (Butcher 1987; Hairer *et al.* 1993; Sanz-Serna & Calvo 1994). We impose order greater than or equal to $r$ by demanding that the expansions of $\psi_h$ and $\phi_h$ differ in $O(h^{r+1})$ terms; this leads to an order condition per rooted tree of order less than or equal to $r$. Note that the expansion of $\psi_h$ is bypassed in the standard Lie methodology outlined above, which works instead with the expansion of the modified field $\tilde{f}_h$; with a Lie-group–Lie-algebra terminology, we may say that here we work in a Lie group of mappings and the standard methodology operates in the corresponding Lie algebra. The rooted trees needed here differ from those used for RK methods in that here, associated with each vertex, there is a positive integer called the type of the vertex. The expansions of $\psi_h$ and $\phi_h$ and the

order conditions are presented in § 3; § 2 contains some preparations. We emphasize that, in the present approach, the order conditions can be written down very easily, as their structure is a mere transcription of the structure of the corresponding rooted trees.

Unfortunately, the order conditions obtained in this way are not all independent; some of them are fulfilled as a consequence of the fulfilment of others (a similar situation arises in Owren & Marthinsen (1997)). In § 4 we identify subsets $\mathcal{H}$, called *Hall* subsets, of the set of all rooted trees such that the order conditions for the members of $\mathcal{H}$ imply all the remaining order conditions. Therefore, in practice, one needs to construct only the rooted trees of a Hall set up to the desired order $r$ (which is an easy job) and then write the associated order conditions (which, as pointed out before, is a matter of transcribing the structure of the corresponding graphs). We show that this procedure is optimal in the sense that the number of order conditions is the same as that one would obtain by following the standard Lie methodology. Some of the proofs for the results in § 4 are rather difficult and require ad hoc notation; we have provided those in isolation in § 5. That section may be skipped, as the final sections do not depend on it.

## 2. Abstract framework

The general form of the expansion of $\psi_h$ in (1.6) can be guessed after explicitly computing the first few terms. We have preferred an alternative approach, introduced in Murua (1999) for RK and related methods. In this section we review the framework in Murua (1999); this will be put to use in the next section.

### (*a*) *Abstract rooted trees*

We consider a countable set $\mathcal{T}$ so that associated with each $u \in \mathcal{T}$ there is an integer $\rho(u)$ and a mapping $F(u) : \mathbb{R}^D \to \mathbb{R}^D$. We say that each $u \in \mathcal{T}$ is a *rooted tree* and that $\rho(u)$ and $F(u)$ are, respectively, *the order* and the *elementary differential* of $u$. At this stage there is no link between the 'abstract' rooted trees used here and the rooted trees studied in graph theory. For simplicity (but this is not necessary (Murua 1999)) we assume throughout that $\rho(u) \geqslant 1$ for all $u \in \mathcal{T}$. It is also assumed that for each integer $\ell \geqslant 1$, the set $\mathcal{T}_\ell = \{u \in \mathcal{T} : \rho(u) = \ell\}$ is finite.

With each mapping $c : \mathcal{T} \to \mathbb{R}$, we associate the formal series of powers of the variable $h$:

$$S(c) = \mathrm{id} + \sum_{u \in \mathcal{T}} h^{\rho(u)} c(u) F(u) = \mathrm{id} + \sum_{\ell=1}^{\infty} h^\ell \sum_{u \in \mathcal{T}_\ell} c(u) F(u). \tag{2.1}$$

### (*b*) *Abstract forests*

By definition, a *forest* $\hat{w}$ (with $m$ rooted trees) is an unordered $m$-tuple of rooted trees $\hat{w} = [u_1, \ldots, u_m]$, $u_i \in \mathcal{T}$, $i = 1, \ldots, m$. An empty forest $[\emptyset]$ having $m = 0$ trees is also considered. The *order* of $\hat{w} = [u_1, \ldots, u_m]$ (not to be confused with $m$) is defined by $\rho(\hat{w}) = \sum_i \rho(u_i)$ and we set $\rho([\emptyset]) = 0$. The set of all forests is denoted by $\hat{\mathcal{T}}$ and, for $\ell = 0, 1, 2, \ldots$, $\hat{\mathcal{T}}_\ell$ represents the subset of $\hat{\mathcal{T}}$ consisting of forests of order $\ell$. While each rooted tree $u$ gives rise to an elementary differential $F(u)$ (a mapping), with each forest $\hat{w} \in \hat{\mathcal{T}}$ we associate an *elementary differential operator*

$X(\hat{w})$ that acts on each mapping $\Phi : \mathbb{R}^D \to V$ ($V$ a finite-dimensional vector space) to give a mapping $X(\hat{w})\Phi : \mathbb{R}^D \to V$. The definition of $X(\hat{w})\Phi$ is as follows. For $\hat{w} = [\emptyset]$, $X(\hat{w})\Phi = \Phi$. For $\hat{w} = [u_1, \dots, u_m]$, $m \geqslant 1$, $y \in \mathbb{R}^D$,

$$(X(\hat{w})\Phi)(y) = \frac{1}{\prod_{i=1}^{m^*} \mu_i!} \Phi^{(m)}(y)(F(u_1)(y), \dots, F(u_m)(y)).$$

Here $\Phi^{(m)}(y)$ is the $m$th Fréchet derivative of $\Phi$ evaluated at $y$; this derivative is an operator that acts on the vectors $F(u_i)(y) \in \mathbb{R}^D$, $i = 1, \dots, m$, to yield the element

$$\Phi^{(m)}(y)(F(u_1)(y), \dots, F(u_m)(y)) \in V.$$

The integer $m^*$, $1 \leqslant m^* \leqslant m$ is the number of distinct elements $u_i^*$ among the $u_j$ in $[u_1, \dots, u_m]$ and, for $i = 1, \dots, m^*$, $\mu_i$ is the number of $u_j$ that are equal to $u_i^*$.

With each mapping $d : \hat{\mathcal{T}} \to \mathbb{R}$, we associate the formal series of differential operators (cf. (2.1))

$$\hat{S}(d) = \sum_{\hat{w} \in \hat{\mathcal{T}}} h^{\rho(\hat{w})} d(\hat{w}) X(\hat{w}) = \sum_{\ell=0}^{\infty} h^{\ell} \sum_{\hat{w} \in \hat{\mathcal{T}}_\ell} d(\hat{w}) X(\hat{w}).$$

### (*c*) *Substitution of a series of elementary differentials in a function*

With the preceding terminology, if $\Phi : \mathbb{R}^D \to V$ is smooth, then the expansion in powers of $h$ of the composition $\Phi \circ S(c)$ is given by (Murua 1999)

$$\Phi \circ S(c) = \hat{S}(c')\Phi, \tag{2.2}$$

where the mapping $c' : \hat{\mathcal{T}} \to \mathbb{R}$ is derived from the mapping $c : \mathcal{T} \to \mathbb{R}$ by the rule

$$c'(\hat{w}) = \Pi_{i=1}^{m} c(u_i), \quad \hat{w} = [u_1, \dots, u_m], \quad c'([\emptyset]) = 1. \tag{2.3}$$

More precisely, (2.2) means that, for $y \in \mathbb{R}^D$,

$$\Phi\left(y + \sum_{u \in \mathcal{T}} h^{\rho(u)} c(u) F(u)(y)\right) = \sum_{\hat{w} \in \hat{\mathcal{T}}} h^{\rho(\hat{w})} c'(\hat{w}) (X(\hat{w})\Phi)(y).$$

## 3. Order conditions

### (*a*) *Preliminaries*

We now look for the series expansion of the result of one step of the method (1.6) as a series $S(c)$ of the form (2.1). It is convenient to seek simultaneously the expansion of (1.6) and the expansions of the intermediate mappings

$$\chi_h^{(k)} = \psi_{\beta_k h}^{[\mathrm{B}]} \circ \psi_{\alpha_k h}^{[\mathrm{B}]*} \circ \cdots \circ \psi_{\beta_2 h}^{[\mathrm{B}]} \circ \psi_{\alpha_2 h}^{[\mathrm{B}]*} \circ \psi_{\beta_1 h}^{[\mathrm{B}]} \circ \psi_{\alpha_1 h}^{[\mathrm{B}]*}, \quad k = 1, 2, \dots, s$$

$$\chi_h^{(k+1/2)} = \psi_{\alpha_{k+1} h}^{[\mathrm{B}]*} \psi_{\beta_k h}^{[\mathrm{B}]} \circ \psi_{\alpha_k h}^{[\mathrm{B}]*} \circ \cdots \circ \psi_{\beta_2 h}^{[\mathrm{B}]} \circ \psi_{\alpha_2 h}^{[\mathrm{B}]*} \circ \psi_{\beta_1 h}^{[\mathrm{B}]} \circ \psi_{\alpha_1 h}^{[\mathrm{B}]*}, \quad k = 0, 2, \dots, s-1.$$

We also set $\chi_h^{(0)} = \mathrm{id}$ and note that $\chi_h^{(s)} = \psi_h$. The expansions of $\chi_h^{(k)}$, $\chi_h^{(k+1/2)}$, $k = 0, 1, \dots, s-1$ are also assumed to be given by series $S(c_k)$, $S(c_{k+1/2})$ *using the same set of rooted trees* $\mathcal{T}$ *and the same elementary differentials used in the series* $S(c)$ *for* $\psi_h$. Our task is to define $\mathcal{T}$, $F(u)$, $c(u)$ so as to satisfy the following requirements.

(1) The set $\mathcal{T}$ is universal, i.e. independent of the differential equation (1.1) being integrated and of the choices of basic method $\psi_h^{[B]}$ and coefficients $\alpha_i$, $\beta_i$ in (1.6).

(2) The elementary differentials $F(u)$ depend on the differential equation (1.1) and on the basic method, but do not depend on the method coefficients $\alpha_i$, $\beta_i$.

(3) The series coefficients $c(u)$ are functions of the method coefficients $\alpha_i$, $\beta_i$, but are independent of the differential equation and of the basic method.

The definitions of $\chi_h^{(k)}$ and $\chi_h^{(k+1/2)}$ imply

$$S(c_{k+1}) = \psi^{[B]}_{\beta_{k+1}h} \circ S(c_{k+1/2}), \quad k = 0, 1, \ldots, s-1, \tag{3.1}$$

so that, if

$$\psi_h^{[B]}(y) = y + \sum_{i=1}^{\infty} h^i f^{(i)}(y) \tag{3.2}$$

is the expansion of the basic method ($f^{(1)} = f$ for consistency), then

$$S(c_{k+1}) = \left(\mathrm{id} + \sum_{i=1}^{\infty} h^i \beta_{k+1}^i f^{(i)}\right) \circ S(c_{k+1/2}), \quad k = 0, 1, \ldots, s-1,$$

or

$$S(c_{k+1}) - S(c_{k+1/2}) = \sum_{i=1}^{\infty} h^i \beta_{k+1}^i f^{(i)} \circ S(c_{k+1/2}).$$

We now resort to (2.2) to write

$$\begin{aligned} S(c_{k+1}) - S(c_{k+1/2}) &= \sum_{i=1}^{\infty} h^i \beta_{k+1}^i \sum_{j=0}^{\infty} h^j \sum_{\hat{w}\in\hat{\mathcal{T}}_j} c'_{k+1/2}(\hat{w}) X(\hat{w}) f^{(i)} \\ &= \sum_{\ell=1}^{\infty} h^\ell \sum_{i=1}^{\ell} \sum_{\hat{w}\in\hat{\mathcal{T}}_{\ell-i}} \beta_{k+1}^i c'_{k+1/2}(\hat{w}) X(\hat{w}) f^{(i)}, \end{aligned}$$

and finally use the definition of $S$ in (2.1) to conclude that

$$\sum_{\ell=1}^{\infty} h^\ell \sum_{u\in\mathcal{T}_\ell} (c_{k+1}(u) - c_{k+1/2}(u)) F(u) = \sum_{\ell=1}^{\infty} h^\ell \sum_{i=1}^{\ell} \sum_{\hat{w}\in\hat{\mathcal{T}}_{\ell-i}} \beta_{k+1}^i c'_{k+1/2}(\hat{w}) X(\hat{w}) f^{(i)}. \tag{3.3}$$

Using this equality we see that, in order to fulfil the three requirements outlined previously, it is plausible to try to ensure that the following three conditions hold.

(i) There is a one-to-one and onto correspondence between $\mathcal{T}_\ell$ and the set of pairs $(i, \hat{w})$, where $i$ is an integer $1 \leqslant i \leqslant \ell$ and $\hat{w}$ is a forest of order $\ell - i$. This condition ensures that in both sides of (3.3) the inner summations comprise the same set of indices.

(ii) The elementary differential associated with $u = (i, \hat{w})$ satisfies

$$F(u) = X(\hat{w}) f^{(i)}. \tag{3.4}$$

(iii) The coefficients $c_{k+1}$ and $c_{k+1/2}$ satisfy, for each $u = (i, \hat{w})$,

$$c_{k+1}(u) - c_{k+1/2}(u) = \beta_{k+1}^i c'_{k+1/2}(\hat{w}), \quad k = 0, 1, \ldots, s-1. \tag{3.5}$$

In the remainder of this section we show that these three conditions can actually be enforced.

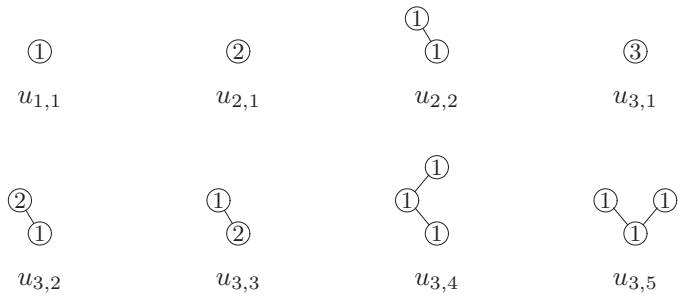


Figure 1. Rooted $\infty$-trees of order up to 3.

### (*b*) *Rooted trees*

To satisfy the condition (i) above, we *define* a rooted tree of order $\ell$ as a pair $(i, \hat{w})$, where $i$ is an integer $1 \leqslant i \leqslant \ell$ and $\hat{w}$ is a forest of order $\ell - i$. This is a recursive definition because, if $\ell - i > 0$, then, by definition, $\hat{w} \in \hat{\mathcal{T}}_{\ell-i}$ is in turn an ordered set of rooted trees $u_i$ having order $\leqslant \ell - i$.

The unique forest $[\emptyset]$ of order $\rho = 0$, gives rise to the unique rooted tree $u_{1,1} = (1, [\emptyset])$ of order $\rho = 1$. There is then a unique forest of order 1, namely $[u_{1,1}]$. At $\rho = 2$, we find two rooted trees $u_{2,1} = (2, [\emptyset])$, $u_{2,2} = (1, [u_{1,1}])$ and therefore three forests $[u_{2,1}]$, $[u_{2,2}]$, $[u_{1,1}, u_{1,1}]$. Then there are five rooted trees with $\rho = 3$, namely $u_{3,1} = (3, [\emptyset])$, $u_{3,2} = (2, [u_{1,1}])$, $u_{3,3} = (1, [u_{2,1}])$, $u_{3,4} = (1, [u_{2,2}])$, $u_{3,5} = (1, [u_{1,1}, u_{1,1}])$, etc.

The rooted trees just introduced are best described in terms of graphs with infinitely many types of vertices. A rooted tree of the form $(i, [\emptyset])$ is described by a single vertex of type $i$ (graphically we draw a circle with the number $i$ inside). The graph that corresponds to the rooted tree $(i, \hat{w})$, $\hat{w} = [u_1, \ldots, u_m]$, $m > 0$, consists of a vertex of type $i$ (the root) joined by edges to the roots of the graphs of the rooted trees $u_i$. The graphs of the rooted trees with $\rho \leqslant 3$ are given in figure 1. Note that the *order of a rooted tree is the sum of the types of its vertices*. Of course it is the graphical interpretation just outlined which justifies the terminology 'rooted tree' that we have been using all along.

Rooted trees whose vertices are all of one and the same type are standard in the theory of RK methods (Butcher 1987; Hairer *et al.* 1993; Sanz-Serna & Calvo 1994). RK–Nyström methods require rooted trees with two types of vertices (Hairer *et al.* 1993; Sanz-Serna & Calvo 1994), and additive RK methods with $N$ parts bring in rooted trees with $N$ types of vertices, often called rooted $N$-trees (Araújo *et al.* 1997). Along the same lines, we will use the terminology *rooted* $\infty$*-trees* to refer to the rooted trees defined above. It will be useful later to note that the set $T$ of the standard RK rooted trees can be identified with the subset of $\infty T$ consisting of those rooted trees whose vertices are all of type 1. This identification preserves the order: if $u \in T$ then its RK order is by definition the number of vertices it has, which coincides with the sum of types $\rho(u)$ because of vertices which have type 1.

If $u = (i, \hat{w}) \in \infty T$, we write $i = \mathrm{tp}(u)$, $\hat{w} = \mathrm{op}(u)$ and say that $i$ and $\hat{w}$ are respectively the *type* and the (*forest of*) *operands* of $u$. The number of vertices in $u$ is denoted by $\nu(u)$.

### (c) Elementary differentials and $\infty B$-series

To satisfy (3.4), we *define*, for $u \in \infty T$,

$$F(u) = X(\mathrm{op}(u))f^{(\mathrm{tp}(u))}. \tag{3.6}$$

This is again a definition recursive with respect to $\rho(u)$ because $X(\mathrm{op}(u))$ involves the elementary differentials of the rooted trees in the forest $\mathrm{op}(u)$.

For the rooted trees in Figure 1, the elementary differentials are successively found to be (primes denote Fréchet derivatives)

$$\begin{gathered} F(u_{1,1}) = f^{(1)}, \quad F(u_{2,1}) = f^{(2)}, \qquad F(u_{2,2}) = f^{(1)\prime}f^{(1)}, \\ F(u_{3,1}) = f^{(3)}, \quad F(u_{3,2}) = f^{(2)\prime}f^{(1)}, \quad F(u_{3,3}) = f^{(1)\prime}f^{(2)}, \\ F(u_{3,4}) = f^{(1)\prime}f^{(1)\prime}f^{(1)}, \quad F(u_{3,5}) = \tfrac{1}{2}f^{(1)\prime\prime}(f^{(1)}, f^{(1)}). \end{gathered}$$

The structure of $F(u)$ mimics that of the graph of $u$ in a way that is familiar from the theory of RK methods; a vertex of type $i$ with $m$ children brings in the $m$th derivative of $f^{(i)}$ acting on $m$ vectors. In particular, if $u \in T$, then the elementary differential defined here coincides (up to the normalization to be discussed presently) with the one it has in an RK context (Butcher 1987; Hairer *et al.* 1993; Sanz-Serna & Calvo 1994).

Note that $F(u_{3,5})$ includes a factor $\frac{1}{2}$. This is because $u_{3,5}$ has a non-trivial set of symmetries; the graph is invariant both by the identity transformation and by the transformation that swaps the two children of the root. In general, $F(u)$ includes a normalizing factor $1/\sigma(u)$, where $\sigma(u)$ is the number of symmetries of $u$ (this factor simplifies many formulae (cf. Butcher & Sanz-Serna 1996).

Each series of the form (2.1) with a set of indices $\mathcal{T} = \infty T$ and elementary differentials (3.6) will be called an $\infty B$-series, thus generalizing the concept of $B$-series introduced by Hairer & Wanner (1974). Note that each $B$-series can be identified with an $\infty B$-series where all coefficients of the rooted $\infty$-trees in $\infty T - T$ vanish.

### (d) Elementary weights

Our definitions of rooted trees and elementary differentials and (3.3) clearly imply (3.5), which can now be rewritten as

$$c_{k+1}(u) - c_{k+1/2}(u) = \beta_{k+1}^{\mathrm{tp}(u)} c'_{k+1/2}(\mathrm{op}(u)), \quad k = 0, 1, \ldots, s-1 \tag{3.7}$$

($c'$ is found via (2.3)).

Thanks to (3.7) we can compute $c_{k+1}(u)$ when $c_{k+1/2}(v)$ is known for rooted trees with $\rho(v) \leqslant \rho(u)$. To find $c_{k+1/2}$ from $c_k$, note that, by definition of $\chi_h^{(k+1/2)}$, $\chi_h^{(k)}$,

$$S(c_{k+1/2}) = \psi^{[\mathrm{B}]*}_{\alpha_{k+1}h} \circ S(c_k). \tag{3.8}$$

We cannot directly proceed as in the derivation of (3.3) because, distinct from the transition from (1.7) to (1.8), the expansion of $\psi_h^{[\mathrm{B}]*}$ cannot be easily retrieved from (3.2) . We rather note that, from (3.8),

$$\psi^{[\mathrm{B}]}_{-\alpha_{k+1}h} \circ S(c_{k+1/2}) = S(c_k),$$

and, by mimicking the steps between (3.1) and (3.3), we conclude that

$$c_{k+1/2}(u) - c_k(u) = -(-\alpha_{k+1})^{\mathrm{tp}(u)} c'_{k+1/2}(\mathrm{op}(u)), \quad k = 0, 1, \ldots, s-1. \tag{3.9}$$

This formula and (3.7), along with the initial condition $c_0(u) = 0$ (recall that $\chi_h^{(0)} = \mathrm{id}$) allow the recursive computation of the coefficients $c(u) = c_s(u)$ in the expansion of $\psi_h$; these coefficients are called the *elementary weights* of the method.

For the rooted trees in figure 1, the elementary weights are easily found to be

$$
\begin{aligned}
c(u_{1,1}) &= \sum_{j=1}^{s}(\alpha_j + \beta_j),\\
c(u_{2,1}) &= \sum_{j=1}^{s}(-\alpha_j^2 + \beta_j^2),\\
c(u_{2,2}) &= \sum_{j=1}^{s}(\alpha_j + \beta_j)\sum_{\ell=1}^{j*}(\alpha_\ell + \beta_\ell),\\
c(u_{3,1}) &= \sum_{j=1}^{s}(\alpha_j^3 + \beta_j^3),\\
c(u_{3,2}) &= \sum_{j=1}^{s}(-\alpha_j^2 + \beta_j^2)\sum_{\ell=1}^{j*}(\alpha_\ell + \beta_\ell),\\
c(u_{3,3}) &= \sum_{j=1}^{s}(\alpha_j + \beta_j)\sum_{\ell=1}^{j*}(-\alpha_j^2 + \beta_j^2),\\
c(u_{3,4}) &= \sum_{j=1}^{s}(\alpha_j + \beta_j)\sum_{\ell=1}^{j*}(\alpha_\ell + \beta_\ell)\sum_{m=1}^{\ell*}(\alpha_m + \beta_m),\\
c(u_{3,5}) &= \sum_{j=1}^{s}(\alpha_j + \beta_j)\sum_{\ell=1}^{j*}(\alpha_\ell + \beta_\ell)\sum_{m=1}^{j*}(\alpha_m + \beta_m)\\
&= \sum_{j=1}^{s}(\alpha_j + \beta_j)\Bigl(\sum_{\ell=1}^{j*}(\alpha_\ell + \beta_\ell)\Bigr)^2.
\end{aligned}
$$

In $c(u_{2,2})$ or $c(u_{3,2})$, the summation with the set of indices $\ell = 1, \ldots, j*$ means that, at $\ell = j$, we use the term $\alpha_\ell$ rather than $\alpha_\ell + \beta_\ell$. Similarly, in $c(u_{3,3})$ we use, at $\ell = j$, the term $-\alpha_\ell^2$ rather than $-\alpha_\ell^2 + \beta_\ell^2$, etc.

It is easy to realize that the expressions for the elementary weights *reflect the structure of the corresponding rooted trees*. Each vertex brings in a summation. The summation corresponding to the root runs from 1 to $s$; the summation corresponding to a vertex $V$ other than the root is 'starred' and runs from 1 to the value of the summation index of the parent of $V$. Types are translated into powers and an even power of $\alpha_j$ must be accompanied by a minus sign.

### (*e*) *The expansion of the true solution and of the local error*

It is well known from the theory of RK methods (see Murua (1999) for a derivation within the present framework) that the expansion of the true flow $\phi_h$ is given by a

$B$-series

$$\mathrm{id} + \sum_{u \in T} h^{\rho(u)} \frac{1}{\gamma(u)} F(u), \tag{3.10}$$

where the set of indices $T$ is the set of (RK) rooted trees and, for $u \in T$, the *density* $\gamma(u)$ is defined by

$$\gamma(u) = \rho(u) \Pi_{v \in \mathrm{op}(u)} \gamma(v)$$

(with the standard convention that if $\mathrm{op}(u)$ is empty, then the product is 1). As noted before, the series (3.10) can be seen as the $\infty B$-series $S(\lambda)$ whose coefficients are given by

$$\left.\begin{aligned} \lambda(u) &= \frac{1}{\gamma(u)}, \quad u \in T, \\ \lambda(u) &= 0, \qquad u \in \infty T - T. \end{aligned}\right\} \tag{3.11}$$

Then the local error $\psi_h - \phi_h$ is the $\infty B$-series with coefficients $c - \lambda$, with $c$ the sequence of elementary weights we computed in the preceding subsection. In particular we obtain the following important result.

**Theorem 3.1.** *The method (1.6) is of order greater than or equal to $r$ for all equations (1.1) and all choices of consistent basic method if, for all rooted trees $u \in \infty T$ with $\rho(u) \leqslant r$, the elementary weight $c(u)$ coincides with the weight $\lambda(u)$ of the true solution given in (3.11).*

For instance, the following eight conditions ensure order greater than or equal to 3: $c(u_{1,1}) = 1$, $c(u_{2,1}) = 0$, $c(u_{2,2}) = \frac{1}{2}$, $c(u_{3,1}) = 0$, $c(u_{3,2}) = 0$, $c(u_{3,3}) = 0$, $c(u_{3,4}) = \frac{1}{6}$, $c(u_{3,5}) = \frac{1}{3}$ (the expressions for the $c$ in terms of the $\alpha_j$ and $\beta_j$ were given above).

For the sake of brevity, we do not discuss here in detail whether the condition $\rho(u) \leqslant r \Rightarrow c(u) = \lambda(u)$ in theorem 3.1 is also necessary for the method to have order greater than or equal to $r$. Roughly speaking this condition is necessary if order greater than or equal to $r$ is demanded for all systems (1.1) and all choices of the basic method. If, however, the basic method is not 'general', then the conditions should not be expected to be necessary. For instance if the basic method is assumed to have order $R > 1$, then $f^{(2)}, \ldots, f^{(R)}$ in the expansion (3.2) coincide with the coefficients of $h^2, \ldots, h^R$ in the expansion of the true flow and they are therefore expressible in terms of $f$; as a result not all elementary differentials are independent.

## 4. Independent order conditions

In the preceding section we have seen that each $u \in \infty T$ introduces an order condition for (1.6); the cardinality of $\infty T_\ell$ grows wildly with $\ell$ and already at $\ell = 3$ there are five rooted trees in $\infty T_\ell$. It is then fortunate that the order conditions are not independent. For instance, from the expressions for the elementary weights in terms of the $\alpha_i$ and $\beta_i$, it is a trivial matter to check that $c(u_{1,1})^2 = 2c(u_{2,2}) + c(u_{2,1})$; thus the order condition $c(u_{2,2}) = \frac{1}{2}$ is a consequence of the order conditions $c(u_{1,1}) = 1$ and $c(u_{2,1}) = 0$. Our task in this section is to determine the relations among the elementary weights of (1.6) and to identify a subset of $\infty T$ that yields a set of mutually independent order conditions.

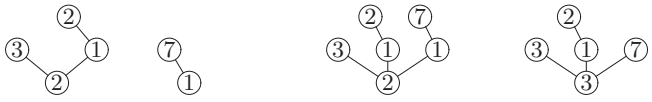


Figure 2. Butcher product and merging product.

### (*a*) *Relations among the elementary weights*

To study the relations that bind the values of the elementary weights we need some operations in $\infty T$. If $u, v \in \infty T$ we define their *Butcher product* $u \cdot v$ by the relations $\mathrm{tp}(u \cdot v) = \mathrm{tp}(u)$, $\mathrm{op}(u \cdot v) = \mathrm{op}(u) \cup \{v\}$; this definition extends that originally suggested by Butcher (1987) for rooted trees in $T$. The *merging product* $u \times v$ is defined by $\mathrm{tp}(u \times v) = \mathrm{tp}(u) + \mathrm{tp}(v)$, $\mathrm{op}(u \times v) = \mathrm{op}(u) \cup \mathrm{op}(v)$. The graphical meaning of these operations is given in figure 2.

The merging product is commutative and associative. The Butcher product is neither, but possesses the property that, for all $u$, $v$, $w$,

$$(u \cdot v) \cdot w = (u \cdot w) \cdot v. \tag{4.1}$$

It will be useful later to observe that, for any $u$, $v$,

$$\rho(u \cdot v) = \rho(u \times v) = \rho(u) + \rho(v),$$
$$\nu(u \cdot v) = \nu(u \times v) + 1 = \nu(u) + \nu(v).$$

After these definitions we have the following result.

**Theorem 4.1.** *Assume that an $\infty B$-series $S(c)$ is the expansion of a composition one-step method (1.6). Then*

$$\forall u, v \in \infty T, \quad c(u \cdot v) + c(v \cdot u) - c(u)c(v) + c(u \times v) = 0, \tag{4.2}$$
$$\forall u, v, w \in \infty T, \quad c(w \cdot (u \cdot v)) + c(w \cdot (v \cdot u)) - c((w \cdot u) \cdot v) + c(w \cdot (u \times v)) = 0. \tag{4.3}$$

*Proof*. We prove (4.2); the proof of (4.3) is similar.

We begin by noticing that it is enough to show that if the coefficients of two $\infty B$-series $S(a)$, $S(b)$ are related by

$$a(u) = b(u) + \alpha^{\mathrm{tp}(u)} b'(\mathrm{op}(u)), \tag{4.4}$$

where $\alpha$ is a real constant and $u$ ranges over all possible rooted $\infty$-trees, then, for each $u, v \in \infty T$,

$$a(u \cdot v) + a(v \cdot u) - a(u)a(v) + a(u \times v) = b(u \cdot v) + b(v \cdot u) - b(u)b(v) + b(u \times v). \tag{4.5}$$

Indeed, the coefficients of the $B$-series of the identity map vanish identically and trivially satisfy the relations (4.2); the coefficients for the $\infty B$-series for the method (1.6) are, in view of (3.7) and (3.9), linked to the coefficients of the identity by a chain of relations of the form (4.4).

Let us then prove (4.5). By using (4.4), the definition (2.3) of the coefficient operation $'$, and the definitions of the products $\cdot$ and $\times$, the left hand-side of (4.5) can

be rewritten as

$$\begin{aligned} b(u\cdot v)+\alpha^{\mathrm{tp}(u)}b'(\mathrm{op}(u))b(v)+b(v\cdot u)+\alpha^{\mathrm{tp}(v)}b'(\mathrm{op}(v))b(u)\\ -(b(u)+\alpha^{\mathrm{tp}(u)}b'(\mathrm{op}(u)))(b(v)+\alpha^{\mathrm{tp}(v)}b'(\mathrm{op}(v)))\\ +b(u\times v)+\alpha^{\mathrm{tp}(u)+\mathrm{tp}(v)}b'(\mathrm{op}(u))b'(\mathrm{op}(v)), \end{aligned}$$

an expression that, after cancellation, may be seen to be equal to the right-hand side of (4.5). ■

The constraint (4.2) is reminiscent of a similar one, first discovered in Sanz-Serna & Abia (1991), in the theory of symplectic RK methods (Sanz-Serna & Calvo 1994). Constraints similar to (4.3) appear in Calvo & Hairer (1995) in the study of some particular families of composition methods.

The following theorem is needed later.

**Theorem 4.2.** *The coefficients $\lambda(u)$ of the $\infty B$-series that corresponds to the true flow mapping $\phi_h$ satisfy the constraints (4.2), (4.3).*

*Proof*. A direct proof using the definition of the coefficients $\lambda$ is straightforward but long (see the similar proof in Sanz-Serna & Abia (1991)). A shorter indirect proof is also possible. For any $r$ there is a method of the form (1.6) of order greater than or equal to $r$ (Sanz-Serna & Calvo 1994, § 13.1). Hence, up to any order $r$, the values $\lambda(u)$ coincide with the values $c(u)$ of a composition method and by the preceding theorem must satisfy the constraints (4.2), (4.3). ■

### (*b*) *Independence of the location of the root*

It is easy to see how to use (4.2) to reduce the number of order conditions. The process is virtually identical to that used in the study of the order conditions of symplectic RK methods (Sanz-Serna & Calvo 1994).

From (4.2) and the corresponding relation for the coefficients $\lambda$ of the true flow, we see that the order conditions for the rooted $\infty$-trees $u\cdot v$ and $v\cdot u$ are equivalent if the order conditions for $u$, $v$ and $u\times v$ hold. We also note that $u$ and $v$ have lower order than $u\cdot v$, $v\cdot u$, while $u\times v$ has fewer vertices than $u\cdot v$, $v\cdot u$. Therefore, if we write the order conditions taking first rooted $\infty$-trees with lower $\rho$ and for each value of $\rho$ we put first the graphs with fewer vertices, then, by the time we deal with $u\cdot v$, $v\cdot u$, the order conditions for $u$, $v$ and $u\times v$ may be assumed to hold. Thus, if we define an equivalence relation in $\infty T$ by declaring each pair $u\cdot v$, $v\cdot u$ to be equivalent, then it is enough to consider an order condition $c=\lambda$ per equivalence class, rather than an order condition per individual rooted $\infty$-tree. These equivalence classes have an obvious graphical interpretation: $v$ and $w$ are equivalent if and only if they possess the same vertices joined by the same edges and just differ in which vertex has been chosen to be the root. For instance, in figure 1, $u_{3,2}$ is equivalent to $u_{3,3}$ and $u_{3,4}$ is equivalent to $u_{3,5}$. Thus each equivalence class corresponds to a *free* (i.e. unrooted) $\infty$-tree, a graph where no vertex has been highlighted to be the root.

Furthermore, by setting $u=v$ in (4.2), we see that the order condition for $u\cdot u$ is a consequence of the order conditions for $u$ and $u\times u$. This shows that we only need to deal with *non-superfluous* free $\infty$-trees, i.e. free $\infty$-trees that do not contain rooted $\infty$-trees of the form $u\cdot u$.

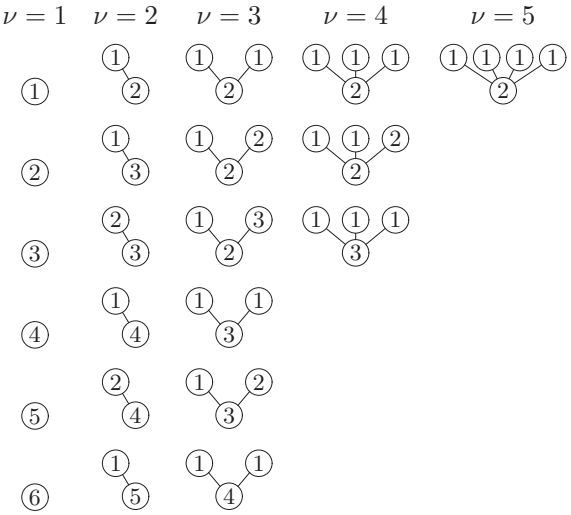


Figure 3. Rooted ∞-trees of order less than or equal to 6 in a Hall set.

After this discussion, the number of conditions for order greater than or equal to 3 has dropped from eight to five ($u_{3,2}$ and $u_{3,3}$ are equivalent, $u_{3,4}$ and $u_{3,5}$ are equivalent and $u_{2,2}$ is superfluous). However, we have still to put to use the constraints (4.3). Dealing with these is more challenging: (4.3) involves *three* rooted ∞-trees (namely $w \cdot (u \cdot v)$, $w \cdot (v \cdot u)$, $(w \cdot u) \cdot v$) with the same value of $\rho$ and $\nu$ and cannot be employed to define an equivalence relation (see the comments at the end of § 2 of Calvo & Hairer (1995)). In the next subsection we deal with (4.2) *and* (4.3) simultaneously and we employ an approach essentially different from that used in the discussion above.

### (*c*) *The main result*

A subset $\mathcal{H} \subset \infty T$ is called *a Hall set* (of rooted ∞-trees) if and only if there is an order relation $<$ in $\mathcal{H}$ in such a way that the following four properties hold:

(i) the order $<$ is total, i.e. if $u, v \in \mathcal{H}$, $u \neq v$, then either $u < v$ or $v < u$;

(ii) the order $<$ is compatible with the function $\nu$ in the sense that $\nu(u) < \nu(v)$ implies $u < v$;

(iii) if $u \in \infty T$ has $\nu(u) = 1$, then $u \in \mathcal{H}$; and

(iv) if $u \in \infty T$ has $\nu(u) > 1$, then $u \in \mathcal{H}$ if and only if $u = v \cdot w$ with $v, w \in \mathcal{H}$, $v > w$.

The construction of Hall sets is discussed in the next subsection. Figure 3 displays, for $\rho \leqslant 6$, the rooted ∞-trees in a Hall set; each graph is $<$ than those on its right and within a column smaller graphs are above larger graphs.

The next theorem is the main result of this paper. The proof is lengthy and technical, and will be given in § 5.

**Theorem 4.3.** *Let $\mathcal{H}$ be a Hall set. The method (1.6) is of order greater than or equal to $r$ for all equations (1.1) and all choices of consistent basic method if, for all rooted trees $u \in \mathcal{H}$ with $\rho(u) \leqslant r$, the elementary weight $c(u)$ coincides with the weight $\lambda(u)$ of the true solution given in (3.11).*

### (*d*) *Generating Hall sets*

In view of properties (i)–(iv) in the corresponding definition, a procedure to generate a Hall set is as follows. We begin by taking all $u \in \infty T$ with $\nu(u) = 1$ and giving them a total order. While there is of course freedom as to how to order these graphs, the order defined by

$$i < j \Rightarrow (i, [\emptyset]) < (j, [\emptyset]) \tag{4.6}$$

seems the obvious choice. In a second step, we obtain all $u \in \mathcal{H}$ with $\nu(u) = 2$ by forming all products $v \cdot w$, $v \in \mathcal{H}$, $w \in \mathcal{H}$, $\nu(v) = \nu(w) = 1$, $v > w$. The graphs obtained in this second step must then be given an order. Again there is freedom as to how to order; a simple way is to have $v_1 \cdot w_1 < v_2 \cdot w_2$ if either $v_1 < v_2$ or $v_1 = v_2$, $w_1 < w_2$. The third step computes all $u \in \mathcal{H}$ with $\nu(u) = 3$ as products $v \cdot w$, where $v, w \in \mathcal{H}$, $\nu(v) = 2$, $\nu(w) = 1$ (which implies $v > w$), etc.

When using this procedure, a point should be taken care of: it is possible that different pairs $v_1$, $w_1$ and $v_2$, $w_2$ give rise to the same product $v_1 \cdot w_1 = v_2 \cdot w_2$. For instance, with a self-explanatory notation, in figure 3 this happens when $v_1 = 3 \cdot 1$, $w_1 = 2$, $v_2 = 3 \cdot 2$, $w_2 = 1$. To avoid such repetitions we impose the rule that, for each $u \in \mathcal{H}$ with $\nu(u) \geqslant 2$, we consider *only* the decomposition $u = v \cdot w$, $v \in \mathcal{H}$, $w \in \mathcal{H}$, $v > w$ that has $w$ as large as possible; this makes sense because, when forming a product, all elements in $\mathcal{H}$ with fewer vertices have already been ordered. In the example above and with the order (4.6), we have $1 < 2$ and therefore the decomposition $v_2 \cdot w_2$, $v_2 = 3 \cdot 2$, $w_2 = 1$ is not allowed; the same rooted $\infty$-tree is obtained as $v_1 \cdot w_1$, $v_1 = 3 \cdot 1$, $w_1 = 2$, where now $w$ is maximal. When this rule is enforced the elements of $\mathcal{H}$ with $\nu \geqslant 3$ are precisely those of the form

$$(u \cdot v) \cdot w, \quad u, v, w, u \cdot v \in \mathcal{H}, \quad u > v, \quad u \cdot v > w, \quad w \geqslant v. \tag{4.7}$$

The requirements $u > v$ and $u \cdot v > w$ arise from property (iv) in the definition of a Hall set. The condition $w \geqslant v$ is imposed because, if $u, v, w, u \cdot v \in \mathcal{H}$, $u > v$, $u \cdot v > w$ and $w < v$, then we would obtain the element $(u \cdot v) \cdot w$ as $(u \cdot w) \cdot v$ (see (4.1) and note that $u \cdot w > v$ because $\nu(u \cdot w) > \nu(u) \geqslant \nu(v)$).

The algorithm in § 3.2 of Murua (1999) may be easily adapted to generate all rooted $\infty$-trees with $\rho \leqslant r$ in a Hall set.

The next theorem counts the number of order conditions by using the Witt formula from the theory of free Lie algebras (Bourbaki 1989). In the formula, $\mu(d)$ denotes the Möbius function: $\mu(1) = 1$, $\mu(d) = (-1)^j$ if $d$ is the product of $j$ distinct primes and $\mu(d) = 0$ otherwise.

**Theorem 4.4.** *A Hall set $\mathcal{H} \subset \infty T$ contains $c(n)$ rooted trees of order $n$, where $c(1) = 1$ and, for $n > 1$,*

$$c(n) = \frac{1}{n} \sum_{d|n} \mu(d) 2^{n/d}.$$

*Proof*. The property (4.7) implies that, in the sense of Bourbaki (1989), $\mathcal{H}$ is a Hall set over the alphabet $1, 2, \ldots$ of all possible types. Then the number of elements of $\mathcal{H}$ with order $n$ equals the dimension of the subspace of elements of degree $n$ of the free Lie algebra generated by arbitrary *symbols* $F^{(1)}, F^{(2)}, \ldots$, where $F^{(i)}$ has degree $i$. For $n > 1$ the dimension of this subspace is given by the Witt formula above, as shown by McLachlan (1995). ■

In particular, $c(2) = 1$, $c(3) = 2$, $c(4) = 3$, $c(5) = 6$, $c(6) = 9$, $c(7) = 18$, $c(8) = 30$, $c(9) = 56$, $c(10) = 99$.

**Corollary 4.5.** *As a corollary we observe that the number of order conditions obtained by our approach is the same as that produced by the standard Lie methodology given in McLachlan (1995); the reason for this is that in McLachlan (1995) one deals with the Lie algebra generated by the functions* $F^{(1)}, F^{(2)}, \ldots$ *in (1.7). This corollary implies that (4.2), (4.3) provide all the constraints that define, within the group of all* $\infty B$*-series, the subgroup of mappings of the form (1.6).*

## 5. Technical proofs

In this section we prove the main theorem 4.3. Throughout the section, $\mathcal{H}$ denotes a Hall set $\subset \infty T$ and we only deal with $\infty B$-series $S(c)$ whose coefficients $c(u)$ satisfy the constraints (4.2), (4.3); recall that this covers the series of the true flow and of any method (1.6).

We say that $u \in \infty T$ *precedes* $v \in \infty T$ if either $\rho(u) < \rho(v)$ or $\rho(u) = \rho(v)$, $\nu(u) < \nu(v)$ (this relation in $\infty T$ should not be confused with the order relation $<$ in $\mathcal{H}$). It is convenient to ignore in the formulae (4.2), (4.3) the coefficients of the rooted $\infty$-trees that precede others in the same formula (this was done in our discussion of (4.2) in § 4 *b*). Thus from (4.2), we write, for each $u$, $v$,

$$c(u \cdot v) \equiv -c(v \cdot u); \tag{5.1}$$

the symbol $\equiv$ is used to relate two expressions $A$, $B$ (namely, $A = c(u{\cdot}v)$, $B = c(v{\cdot}u)$) that differ in a function (namely $c(u)c(v) - c(u \times v)$) that is a polynomial in the coefficients of rooted $\infty$-trees (namely $u$, $v$, $u \times v$) that precede the rooted $\infty$-trees (namely $u \cdot v$, $v \cdot u$) whose coefficients feature in $A$ and $B$.

Similarly, (4.3) implies, for each $u$, $v$,

$$c(w \cdot (u \cdot v)) \equiv -c(w \cdot (v \cdot u)) + c((w \cdot u) \cdot v). \tag{5.2}$$

If $A \equiv B$, then we say that $A$ and $B$ are *equivalent*.

Since the Butcher product is not associative, iterated products require the use of brackets. In order to avoid difficult-to-read expressions with multiple brackets, we introduce the notations ($\ell \geqslant 2$)

$$w_1 \cdot (w_2 \cdot (\cdots (w_{\ell-1} \cdot w_\ell))) = w_1 \circ w_2 \circ \cdots \circ w_\ell, \tag{5.3}$$

$$(((w_1 \cdot w_2) \cdot \cdots) \cdot w_{\ell-1}) \cdot w_\ell = w_1 w_2 \cdots w_\ell. \tag{5.4}$$

Two special cases of (5.4) will be abbreviated further: we write $v^\ell$ rather than $vv \cdots v$ ($\ell$ times, $\ell \geqslant 1$) and $wv^\ell$ rather than $wvv \cdots v$ ($v$ features $\ell$ times as a factor, $\ell \geqslant 1$).

After all these preliminaries we present five lemmas.

**Lemma 5.1.** *For all $\ell \geqslant 1$ and $w_1, w_2, \ldots, w_\ell, u, v \in \circ\!\!\infty T$,*

$$c(w_1 \circ w_2 \circ \cdots \circ w_\ell \circ v) \equiv (-1)^{\ell+1} c((w_\ell \circ w_{\ell-1} \cdots \circ w_1) \circ v), \tag{5.5}$$

$$c(w_1 \circ \cdots \circ w_\ell \circ u \circ v) \equiv -c(w_1 \circ \cdots \circ w_\ell \circ v \circ u) + c(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell \circ v) \circ u). \tag{5.6}$$

*Proof*. We establish (5.5) by induction. There is nothing to prove when $\ell = 1$. For $\ell \geqslant 2$, we successively use the definition of $\circ$, the induction hypothesis, the constraint (5.1) and the property (4.1) in order to write

$$\begin{aligned} c(w_1 \circ \cdots \circ w_\ell \circ v) &= c(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell \circ v)) \\ &\equiv (-1)^{\ell} c((w_{\ell-1} \circ \cdots \circ w_1) \circ (w_\ell \circ v)) \\ &\equiv (-1)^{\ell+1} c((w_\ell \circ v) \circ (w_{\ell-1} \circ \cdots \circ w_1)) \\ &\equiv (-1)^{\ell+1} c((w_\ell \circ w_{\ell-1} \circ \cdots \circ w_1) \circ v). \end{aligned}$$

To prove (5.6), we use the definition of $\circ$, (5.5), (5.2) and (4.1):

$$\begin{aligned} &c(w_1 \circ \cdots \circ w_\ell \circ u \circ v) \\ &\quad = c(w_1 \circ \cdots \circ w_\ell \circ (u \circ v)) \\ &\quad \equiv (-1)^{\ell+1} c((w_\ell \circ \cdots \circ w_1) \circ (u \circ v)) \\ &\quad \equiv (-1)^{\ell} c((w_\ell \circ \cdots \circ w_1) \circ (v \circ u)) + (-1)^{\ell+1} c(((w_\ell \circ \cdots \circ w_1) \circ v) \circ u) \\ &\quad \equiv (-1)^{\ell} c((w_\ell \circ \cdots \circ w_1) \circ (v \circ u)) + (-1)^{\ell+1} c(((w_\ell \circ v) \circ w_{\ell-1} \circ \cdots \circ w_1) \circ u). \end{aligned}$$

We now invoke (5.5) to conclude the proof. ■

**Lemma 5.2.** *For all $\ell, m \geqslant 1$ and $w_1, \ldots, w_\ell$, $v \in \circ\!\!\infty T$,*

$$c(w_1 \circ \cdots \circ w_\ell \circ (v^m)) \equiv \frac{1}{m} c(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell v^m)). \tag{5.7}$$

*Proof*. Induction in $m$. The case $m = 1$ is trivial. For $m > 1$, the definitions of $\circ$ and $v^m$ and (5.6) imply

$$\begin{aligned} &c(w_1 \circ \cdots \circ w_\ell \circ (v^m)) \\ &\quad = c(w_1 \circ \cdots \circ w_\ell \circ (v^{m-1}) \circ v) \\ &\quad \equiv -c(w_1 \circ \cdots \circ w_\ell \circ v \circ (v^{m-1})) + c(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell \circ v) \circ (v^{m-1})). \end{aligned}$$

By the induction hypothesis,

$$\begin{aligned} c(w_1 \circ \cdots \circ w_\ell \circ (v^m)) \equiv &-\frac{1}{m-1} c(w_1 \circ \cdots \circ w_\ell \circ (v v^{m-1})) \\ &+ \frac{1}{m-1} c(w_1 \circ \cdots \circ w_{\ell-1} \circ ((w_\ell \circ v) v^{m-1})). \end{aligned}$$

After noticing that $v v^{m-1} = v^m$ and $(w_\ell \circ v) v^{m-1} = w_\ell v^m$, the last relation leads to the result. ■

**Lemma 5.3.** *For all $m \geqslant 1$ and $v \in \circ\!\!\infty T$,*

$$c(v^{m+1}) \equiv 0. \tag{5.8}$$

*Proof*. From (5.1) and (5.7),

$$c(v^{m+1}) = c(v^m \cdot v) \equiv -c(v \cdot v^m) \equiv -\frac{1}{m}c(vv^m) = -\frac{1}{m}c(v^{m+1}).$$

■

**Lemma 5.4.** *If $\ell \geqslant 1$, $m \geqslant 1$, $w_1, \ldots, w_\ell \in \infty T$, $v_1, \ldots, v_m \in \mathcal{H}$ with $v_1 \leqslant v_2 \leqslant \cdots \leqslant v_m$, then*

$$c(w_1 \circ \cdots \circ w_\ell \circ (v_1 \cdots v_m))$$

*is equivalent to a linear combination of*

$$c(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell v_1 \ldots v_m))$$

*and terms of the form*

$$c(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell u_1 \ldots u_k))$$

*with $k < m$, $u_i \in \mathcal{H}$ and*

$$\sum \rho(u_i) = \sum \rho(v_j), \quad \sum \nu(u_i) = \sum \nu(v_j).$$

*Proof*. Once more we resort to induction in $m$ and once more the case $m = 1$ is trivial. If $m > 1$ and $v_1 = v_2 = \cdots = v_m$, the result is true by (5.7). If $m > 1$ and $v_1 < v_m$, we use (5.6) to write

$$\begin{aligned} c(w_1 \circ \cdots \circ w_\ell \circ (v_1 \cdots v_m)) &= c(w_1 \circ \cdots \circ w_\ell \circ (v_1 \cdots v_{m-1}) \circ v_m) \\ &\equiv -c(w_1 \circ \cdots \circ w_\ell \circ v_m \circ (v_1 \cdots v_{m-1})) \\ &\quad + c(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell \circ v_m) \circ (v_1 \cdots v_{m-1})). \end{aligned}$$

By the induction hypothesis, $c(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell \circ v_m) \circ (v_1 \cdots v_{m-1}))$ is a linear combination like those considered in the lemma. This leaves us with the coefficient $c(w_1 \circ \cdots \circ w_\ell \circ v_m \circ (v_1 \cdots v_{m-1}))$, that, by the induction hypothesis, is equivalent to a linear combination of

$$c(w_1 \circ \cdots \circ w_\ell \circ (v_m v_1 \cdots v_{m-1})) \tag{5.9}$$

and terms of the form

$$c(w_1 \circ \cdots \circ w_\ell \circ (v_m u_1 \cdots u_k)) \tag{5.10}$$

with $k < m - 1$, $u_i \in \mathcal{H}$, $\rho(u_1) + \cdots + \rho(u_k) = \rho(v_1) + \cdots + \rho(v_{m-1})$ and $\nu(u_1) + \cdots + \nu(u_k) = \nu(v_1) + \cdots + \nu(v_{m-1})$. After noting that $v_m v_1 \in \mathcal{H}$, we can once more use the induction hypothesis in (5.9) and (5.10) and this leads to the result. ■

**Lemma 5.5.** *If $m \geqslant 1$, $v_1, v_2, \ldots, v_m \in \mathcal{H}$ with $v_1 \leqslant v_2 \leqslant \cdots \leqslant v_m$, then the coefficient $c(v_1 v_2 \cdots v_m)$ is equivalent to a linear combination of terms of the form $c(u)$ with $u \in \mathcal{H}$, $\rho(u) = \rho(v_1 v_2 \cdots v_m)$, $\nu(u) = \nu(v_1 v_2 \cdots v_m)$.*

*Proof*. We again use induction in $m$. The case $m = 1$ is trivial and the case $m > 1$, $v_1 = \cdots = v_m$ is given by (5.8). For $m > 1$, $v_1 < v_m$, we write

$$c(v_1 v_2 \cdots v_m) = c((v_1 \cdots v_{m-1}) \circ v_m) \equiv -c(v_m \circ (v_1 \cdots v_{m-1})).$$

We apply lemma 5.4 to reduce $c(v_m \circ (v_1 \cdots v_{m-1}))$ to a linear combination and the proof concludes by using an argument similar to that in the proof of lemma 5.4. ■

After the preceding lemmas, we are almost in a position to prove theorem 4.3. The idea of the proof is that lemma 5.4 provides a way to reduce the order condition of any rooted $\infty$-tree $u$ to the order conditions of 'simpler' rooted $\infty$-trees $w_i$. Here 'simpler' means that going from $u$ to the $w_i$ reduces the value of a 'Lyapunov' function $p$ that we define next. For $u \in \infty T$ with $\nu(u) = 1$ we set $p(u) = 0$; for $u \in \infty T$ with $\mathrm{op}(u) = \{w_1, \dots, w_\ell, v_1, \dots, v_m\}$, $w_i \in \infty T - \mathcal{H}$, $v_j \in \mathcal{H}$, we set recursively

$$p(u) = \ell + p(w_1) + \cdots + p(w_\ell).$$

The following two properties are obvious consequences of the definition of $p$:

(P1) if $u \in \infty T$, $v \in \mathcal{H}$, then $p(u \cdot v) = p(u)$;

(P2) if $u \in \infty T$, $v \in \infty T - \mathcal{H}$, then $p(u \cdot v) = 1 + p(u) + p(v)$.

By induction we obtain two additional facts:

(P3) if $w_1, \dots, w_\ell \in \infty T$, $v \in \infty T - \mathcal{H}$, then $p(w_1 \circ w_2 \circ \cdots w_\ell \circ v) = \ell + p(w_1) + \cdots + p(w_\ell) + p(v)$; and

(P4) if $w_1, \dots, w_\ell \in \infty T$, $u_1, \dots, u_k \in \mathcal{H}$, then $p(w_1 \circ \cdots \circ w_{\ell-1} \circ (w_\ell u_1 \cdots u_k)) \leqslant \ell - 1 + p(w_1) + \cdots + p(w_\ell)$.

We need the two further properties, that are easy consequences of the fact that $u \cdot v \in \mathcal{H}$ whenever $u, v \in \mathcal{H}$, $u > v$:

(P5) $p(u) = 0$ if and only if there exist $v_1, \dots, v_m \in \mathcal{H}$, $v_1 \leqslant \cdots \leqslant v_m$ such that $u = v_1 \cdots v_m$; and

(P6) if $p(u) \neq 0$, then there exist $w_1, \dots, w_\ell \in \infty T$, $v \in \infty T - \mathcal{H}$, such that $p(v) = 0$ and $u = w_1 \circ \cdots \circ w_\ell \circ v$.

The next result is the heart of the matter.

**Lemma 5.6.** *For any $u \in \infty T$, $c(u)$ is equivalent to a linear combination of terms of the form $c(v)$, with $v \in \mathcal{H}$, $\rho(v) = \rho(u)$, $\nu(v) = \nu(u)$.*

*Proof*. By induction in the value of $p$. By property (P5), the case $p(u) = 0$ is covered by lemma 5.5. If $p(u) > 1$, then, by (P6), $u = w_1 \circ \cdots w_\ell \circ v$, with $p(v) = 0$, $v \in \infty T - \mathcal{H}$, which, by (P3) leads to

$$p(u) = \ell + p(w_1) + \cdots + p(w_\ell).$$

By lemma 5.4, $c(u)$ is equivalent to a linear combination of terms $c(u^*)$ that, by (P4), have

$$p(u^*) \leqslant \ell - 1 + p(w_1) + \cdots + p(w_\ell) < p(u)$$

and may, by the induction hypothesis, be written as linear combinations. ■

We are finally ready to prove theorem 4.3. We have to show that $c(u) = \lambda(u)$ for each $u \in \infty T$ with $\rho(u) \leqslant r$. We do this by induction on $\rho(u)$. The case $\rho(u) = 1$ is trivial because then $u \in \mathcal{H}$. To go from $\rho - 1$ to $\rho > 2$ we use a new induction, this time in $\nu$. If $\nu = 1$ then there is nothing to be proved because $\mathcal{H}$ contains all rooted $\infty$-trees with one vertex. To go from $\nu - 1$ to $\nu > 2$ vertices we note that, by the last lemma, $c(u)$ is given by a function $F(c(u_1^*), \dots, c(u_\ell^*), c(u_1^{**}), \dots, c(u_m^{**}))$,

where the $u_i^* \in \mathcal{H}$ have $\rho(u_i^*) = \rho(u)$, $\nu(u_i^*) = \nu(u)$ and the $u_j^{**}$ have either $\rho(u_i^{**}) < \rho(u)$, or $\rho(u_i^{**}) = \rho(u)$, $\nu(u_i^{**}) < \nu(u)$. But, due to theorem 4.2, $\lambda(u) = F(\lambda(u_1^*), \ldots, \lambda(u_\ell^*), \lambda(u_1^{**}), \ldots, \lambda(u_m^{**}))$. Then $c(u) = \lambda(u)$, because $c(u_i^*) = \lambda(u_i^*)$ by the hypothesis of the theorem being proved and $c(u_j^{**}) = \lambda(u_j^{**})$ by the induction hypothesis. ■

## 6. The case of the self-adjoint basic method

In this section we discuss composition methods of the format (1.5). In fact these methods are the particular case $\alpha_i = 0$, $i = 1, \ldots, s$, of the format (1.6) we have been discussing. As pointed out in the introduction, there is little point in methods (1.5) if the basic method is 'general': with $\alpha_i = 0$, $i = 1, \ldots, s$, the order conditions $\sum(-\alpha^{2j} + \beta_j^{2j}) = 0$, $j = 1, 2, \ldots$, corresponding to the rooted $\infty$-trees with one vertex and even order cannot possibly be enforced. Therefore from now on we make the assumption that the basic method is *self-adjoint*. Because we are then dealing with a basic method that is not 'general', not all order conditions in theorem 3.1 are necessary. Our task is to find a minimal set of order conditions.

We need the following lemma, where the superscript [C] refers to Cholesky, in analogy with a similar construction familiar in numerical linear algebra (see, for example, Golub & Van Loan 1989).

**Lemma 6.1.** *A method $\psi_{h,f}^{[\mathrm{B}]}$ is self-adjoint if and only if there is a second method $\psi_{h,f}^{[\mathrm{C}]}$ with*

$$\psi_{h,f}^{[\mathrm{B}]} = \psi_{h/2,f}^{[\mathrm{C}]} \circ \psi_{h/2,f}^{[\mathrm{C}]*}.$$

*Proof*. The 'if' part is trivial. For the 'only if' implication, we invoke the modified vector field to write

$$\psi_{h,f}^{[\mathrm{B}]} = \phi_{h,\tilde{f}_h^{[\mathrm{B}]}} = \phi_{h/2,\tilde{f}_h^{[\mathrm{B}]}} \circ \phi_{h/2,\tilde{f}_h^{[\mathrm{B}]}},$$

where the second equality comes from the semigroup property of the flow. We define $\psi_{h/2,f}^{[\mathrm{C}]} = \phi_{h/2,\tilde{f}_h^{[\mathrm{B}]}}$ and to finish the proof we prove that

$$\psi_{h/2,f}^{[\mathrm{C}]*} = \phi_{h/2,\tilde{f}_h^{[\mathrm{B}]}}.$$

By the definition of the adjoint method and the semigroup property of flows, we have that

$$\psi_{h/2,f}^{[\mathrm{C}]*} = (\psi_{-h/2,f}^{[\mathrm{C}]})^{-1} = (\phi_{-h/2,\tilde{f}_{-h}^{[\mathrm{B}]}})^{-1} = \phi_{h/2,\tilde{f}_{-h}^{[\mathrm{B}]}}.$$

The last mapping coincides with $\phi_{h/2,\tilde{f}_h^{[\mathrm{B}]}}$ because for self-adjoint methods the modified vector fields are an even function of $h$ (cf. (1.7), (1.8)). ■

Note that the construction in the proof only yields $\psi_{h,f}^{[\mathrm{C}]}$ as a formal power series. This poses no difficulty in our context where we only need to work modulo $O(h^{r+1})$ terms.

We use the lemma to rewrite the method (1.5) being studied as

$$\psi_h = \psi_{\gamma_s h/2}^{[\mathrm{C}]} \circ \psi_{\gamma_s h/2}^{[\mathrm{C}]*} \circ \cdots \circ \psi_{\gamma_2 h/2}^{[\mathrm{C}]} \circ \psi_{\gamma_2 h/2}^{[\mathrm{C}]*} \circ \psi_{\gamma_1 h/2}^{[\mathrm{C}]} \circ \psi_{\gamma_1 h/2}^{[\mathrm{C}]*} \tag{6.1}$$

and treat (6.1) as a particular instance of the methods (1.6) considered in the preceding sections. Because (6.1) is not the most general case of (1.6), the elementary weights will possess some special properties, as we investigate next.

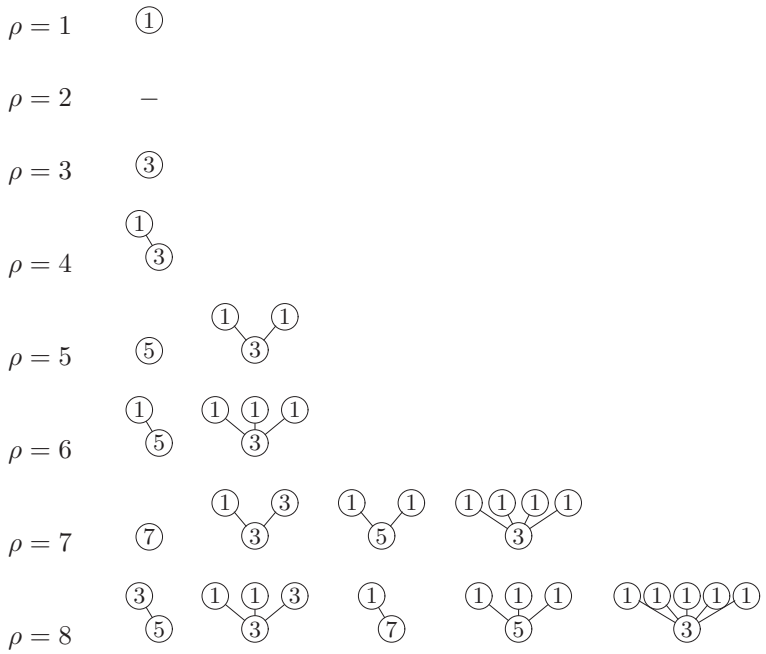


Figure 4. Rooted ∞-trees for compositions of a self-adjoint basic method

**Lemma 6.2.** *If $u \in \infty T$ has even* $\mathrm{tp}(u)$*, then the elementary weight of the method (6.1) vanishes.*

*Proof*. The outermost summation in the expression for $c(u)$ includes the term $-(-\gamma_j)^{\mathrm{tp}(u)} + \gamma_j^{\mathrm{tp}(u)}$. ■

Since $\lambda(u) = 0$ if $\mathrm{tp}(u) = 0$, the lemma guarantees that all the order conditions for the $u$ of even type hold regardless of the choice of the $\gamma_i$.

**Theorem 6.3.** *Let $\mathcal{H}$ be a Hall set. The method (6.1) is of order greater than or equal to $r$ for all equations (1.1) and all choices of a consistent self-adjoint basic method if, for all rooted trees $u \in \mathcal{H}$ with $\rho(u) \leqslant r$ consisting only of vertices of odd type, the elementary weight $c(u)$ coincides with the weight $\lambda(u)$ of the true solution given in (3.11).*

*Proof*. If the method $u \in \mathcal{H}$ has $\rho(u) = r$ and a vertex of even type, then, by the discussion in § 4 *b*, we may assume that that vertex is the root. Then lemma 6.2 shows that the order condition is satisfied and we apply the main theorem 4.3. ■

A set of graphs whose order conditions guarantee order greater than or equal to 8 is given in figure 4; for convenience the graphs have been grouped by order. One may either form a full Hall set as in figure 3 and weed out the graphs with at least one even vertex, or alternatively start with the rooted ∞-trees with $\nu = 1$ and odd type and then recursively form products $u \cdot v$, $u > v$. The number of order conditions is again the same as one would obtain by the standard Lie methodology (McLachlan 1995).

Here is a list of the order conditions for order greater than or equal to 6.

$$\sum_{j=1}^{s}\gamma_j = 1, \quad \sum_{j=1}^{s}\gamma_j^3 = 0, \quad \sum_{j=1}^{s}\gamma_j^3\sum_{\ell=1}^{j*}\gamma_\ell = 0, \quad \sum_{j=1}^{s}\gamma_j^5 = 0,$$
$$\sum_{j=1}^{s}\gamma_j^3\Big(\sum_{\ell=1}^{j*}\gamma_\ell\Big)^2 = 0, \quad \sum_{j=1}^{s}\gamma_j^5\sum_{\ell=1}^{j*}\gamma_\ell = 0, \quad \sum_{j=1}^{s}\gamma_j^3\Big(\sum_{\ell=1}^{j*}\gamma_\ell\Big)^3 = 0.$$

Here summation with the set of indices $\ell = 1, \ldots, j*$ means that at $\ell = j$ we use $\frac{1}{2}\gamma_j$ instead of $\gamma_j$. The left-hand side of each order condition is a simple transcription of the structure of the corresponding graph. It is also useful to compare the simplicity of the order conditions obtained here with the messy unstructured recursions in Yoshida (1990); furthermore the approach in Yoshida (1990) assumes symmetry in the coefficients of (1.5). The expressions in Suzuki (1992) are only slightly more complicated than those found here, but again Suzuki (1992) operates only under symmetry of the coefficients and provides no mnemonics in terms of graphs.

Additional reductions in the number of order conditions are possible if in (1.5) the coefficients are chosen symmetrically $\gamma_i = \gamma_{s-i}$, $i = 1, \ldots, s$. Then the composition method is itself self-adjoint and hence of even order; as a consequence, one has to impose explicitly only the conditions for odd $\rho$.

## 7. Splitting methods

To end the paper, we briefly consider the family of methods (1.4) for split systems (1.1), $f = f_1 + f_2$. As pointed out by McLachlan (1995), the best way of dealing with such methods is to choose the simplest split method (1.2) as a basic method to be used in the format (1.6). This yields a splitting method of the form (1.4), with

$$a_i = \alpha_i + \beta_i, i = 1 \ldots, s, \quad b_0 = \alpha_1, \quad b_i = \alpha_{i+1} + \beta_i, i = 1, \ldots, s-1, \quad b_s = \alpha_s.$$

Note that, when following this approach, there are $2s$ free parameters $\alpha_i$, $\beta_i$ and as a result one does not fully recover the $(2s+1)$-parameter family (1.4): fortunately the methods that are left out are those for which $b_0 + \cdots + b_s \neq a_1 + \cdots + a_s$ and these cannot be consistent!

Along the same lines, it is possible to start by taking the self-adjoint Strang method (1.3) as a basic method in (1.5) and then proceed as in § 6.

## References

Araújo, A. L., Murua, A. & Sanz-Serna, J. M. 1997 Symplectic methods based on decompositions. *SIAM J. Numer. Analysis* **34**, 1926–1947.

Bourbaki, N. 1989 *Lie groups and Lie algebras*, ch. 1–3. New York: Springer.

Butcher, J. C. 1963 Coefficients for the study of Runge–Kutta integration processes. *J. Austral. Math. Soc.* **3**, 185–201.

Butcher, J. C. 1964 On Runge–Kutta processes of high order. *J. Austral. Math. Soc.* **4**, 179–194.

Butcher, J. C. 1987 *The numerical analysis of ordinary differential equations*. Chichester: Wiley.

Butcher, J. C. & Sanz-Serna, J. M. 1996 The number of conditions for a Runge–Kutta method to have effective order *p*. *Appl. Numer. Math.* **22**, 103–111.

Calvo, M. P. & Hairer, E. 1995 Further reduction in the number of independent order conditions for symplectic, explicit partitioned Runge–Kutta and Runge–Kutta–Nyström methods. *Appl. Numer. Math.* **18**, 107–114.

Golub, G. H. & Van Loan, C. F. 1989 *Matrix computations*, 2nd edn. North Oxford Academic.

Hairer, E. & Wanner, G. 1974 On the Butcher group and general multivalue methods. *Computing* **13**, 1–15.

Hairer, E., Nørsett, S. P. & Wanner, G. 1993 *Solving ordinary differential equations. I. Nonstiff problems*, 2nd edn. Berlin: Springer.

McLachlan, R. I. 1995 On the numerical integration of ordinary differential equations by symmetric composition methods. *SIAM J. Sci. Comput.* **16**, 151–168.

Murua, A. 1999 Formal series and numerical integrators. I. Systems of ODEs and symplectic integrators. *Appl. Numer. Math.* **29**, 221–251.

Owren, B. & Marthinsen, A. 1997 Integration methods based on rigid frames. Preprint Numerics no. 1/1997. The Norwegian University of Science and Technology, Trondheim, Norway.

Sanz-Serna, J. M. 1997 Geometric integration. In *The state of the art in numerical analysis* (ed. I. S. Duff & G. A. Watson), pp. 121–143. Oxford: Clarendon.

Sanz-Serna, J. M. & Abia, L. 1991 Order conditions for canonical Runge–Kutta schemes. *SIAM J. Numer. Analysis* **28**, 1081–1096.

Sanz-Serna, J. M. & Calvo, M. P. 1994 *Numerical Hamiltonian problems*. London: Chapman & Hall.

Sanz-Serna, J. M. & Portillo, A. 1996 Classical integrators for wave-packet dynamics. *J. Chem. Phys.* **104**, 2349–2355.

Strang, G. 1968 On the construction and comparison of difference schemes. *SIAM J. Numer. Analysis* **5**, 506–517.

Suzuki, M. 1992 General theory of higher-order decompositions of exponential operators and symplectic integrators. *Phys. Lett.* A **165**, 387–395.

Yoshida, H. 1990 Construction of higher order symplectic integrators. *Phys. Lett.* A **150**, 262–268.