From Runge-Kutta methods to Hopf algebras of rooted trees

A. Murua
Konputazio Zientziak eta A.A. saila, Informatika Fakultatea,
EHU/UPV, Donostia/San Sebastian, Spain

May 27, 2021

Contents

1 Numerical integration methods for ordinary differential equations

1.1 Introduction

We are interested in numerical methods for systems of ordinary differential equations (ODEs) of the form

$$\frac{dy}{dt} = f(y), \quad f : \mathcal{U} \subset \mathbb{R}^d \to \mathbb{R}^d \tag{1.1}$$

where t is the independent variable ("time") and $\mathcal{U}$ is a nonempty open set of $\mathbb{R}^d$. Usually, $\mathcal{U}$ is called the *phase space* of (1.1) and $f : \mathcal{U} \subset \mathbb{R}^d \to \mathbb{R}^d$ is referred to as the vector field. We will assume throughout that the vector field f is smooth in the sense that it has continuous derivatives of all orders.

Recall that the differential equation (1.1) supplemented with the initial condition $y(t_0) = y_0$ has a unique solution $y(t)$. Numerical integration methods aim at obtaining approximations $y_k \approx y(t_k)$ at the time levels $t_1, t_2, \ldots$, with small step sizes $h_k = t_k - t_{k-1}$. If a constant step size $h_k = h$ is considered, then $t_k = kh$ for all k.

In these notes we focus on *one-step methods* to integrate initial-value problems for systems of differential equations (1.1). The simplest (one step) numerical method for (1.1) is the so-called (explicit or forward) Euler method: Consider the one-parameter family of maps $\psi_{h\,f} : \mathbb{R}^d \to \mathbb{R}^d$ defined as

$$\psi_{h\,f}(y) = y + hf(y). \tag{1.2}$$

A sequence $\{y_1, y_2, \ldots\}$ of approximations to the values $y(t_k)$ of the solution of (1.1) supplemented with the initial condition $y(0) = y_0$ at times t_k ($k = 1, 2, \ldots$) is obtained in a step-by-step manner as

$$y_k = \psi_{h_k\,f}(y_{k-1}), \quad k = 1, 2, \ldots \tag{1.3}$$

The smaller the step-size h, the more accurate approximations. This is a consequence of the fact that, for $\tilde{y}(h) := \psi_{hf}(y_0)$, one has that

$$\frac{d}{dh}\tilde{y}(t) = f(\tilde{y}(h)) + hR(y_0, h),$$

where $R(y_0, h) := \frac{1}{h}(f(y_0) - f(y_0 + hf(y_0)))$ is uniformly bounded for sufficiently small $|h|$. That is, $\tilde{y}(t)$ approximately satisfies the original differential equation (1.1) for small values of time $t = h$.

In this sense, a method is of *order* n (n a positive integer), if for each fixed y, there exists $h_0(y), C(y) > 0$ such that, for all $h \in [-h_0(y), h_0(y)]$,

$$||\frac{\partial}{\partial h}\psi_{h\,f}(y) - f(\psi_{h\,f}(y))|| \leq C(y)h^{n+1}. \tag{1.4}$$

From our preceding discussion, it is clear that Euler's method is of order $n = 1$.

Another simple example of a one-step integrator is the explicit trapezoidal method (due to Runge), where

$$\psi_{h\,f}(y) = y + \frac{h}{2}(f(y) + f(y + hf(y))) \tag{1.5}$$

is used in (1.3) instead of (1.2). It is not difficult to check that method (1.5) is of order 2 (that is, (1.4) holds for $n = 2$).

1.2 Runge-Kutta methods

A *Runge-Kutta* (RK) method with s *stages* is specified by a *RK tableau* of real constants

$$\begin{array}{|ccc} a_{11} & \cdots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{s1} & \cdots & a_{ss} \\ \hline b_1 & \cdots & b_s \end{array}\ . \tag{1.6}$$

When applied to the system (1.1), the method corresponding to (1.6) advances the numerical solution from time t_{k-1} to time $t_k = t_{k-1} + h$ through the relation $y_k = \psi_{h\,f}(y_{k-1})$, where

$$\psi_{h\,f}(y) = y + h\sum_{i=1}^{s} b_i f(Y_i), \tag{1.7}$$

and the vectors Y_i (the so-called *internal stages*) are determined by the relations

$$Y_i = y + h\sum_{j=1}^{s} a_{ij} f(Y_j), \quad i = 1, \ldots, s. \tag{1.8}$$

If the matrix A is strictly lower triangular (i.e., if $a_{ij} = 0$ whenever $i \leq j$), the equations (1.8) provide a recursion for explicitly computing each Y_i in terms of the preceding internal stages:

$$\begin{aligned} Y_1 &= y_n, \\ Y_2 &= y_n + h\,a_{21} f(Y_1), \\ Y_3 &= y_n + h\,a_{31} f(Y_1) + h\,a_{32} f(Y_2) \\ &\ \vdots \end{aligned}$$

The method is then called *explicit*. The computation of one step of an explicit Runge-Kutta (RK) method thus requires s evaluations of the function f.

For general matrices A, the method is *implicit,* and (1.8) provides a coupled system of $s \times d$ algebraic equations for the $s \times d$ components of the stage vectors. By the Implicit Function Theorem, given $y_0 \in \mathcal{U}$, there exists $h_0 > 0$ and a neighbourhood $\mathcal{V}$ of $(y_0, \ldots, y_0) \in \mathbb{R}^{s\,d}$ such that the algebraic system (1.8) has in $\mathcal{V}$ a unique solution that smoothly depends on $h \in [-h_0, h_0]$.

For the tableau

$$\begin{array}{|c} 0 \\ \hline 1 \end{array}\,, \tag{1.9}$$

the equation (1.8) reads $Y_1 = y$ and (1.7) is

$$\psi_{h\,f} = y + hf(Y_1) = y + hf(y),$$

which is the explicit Euler formula (1.2). The RK method given by the tableau

$$\begin{array}{|cc} 0 & 0 \\ 1 & 0 \\ \hline 1/2 & 1/2 \end{array}$$

corresponds to the explicit trapezoidal method (1.5).

The two examples of RK tableaux considered so far correspond to explicit RK methods. The RK tableau

$$\begin{array}{|cc} 0 & 0 \\ 1/2 & 1/2 \\ \hline 1/2 & 1/2 \end{array} \tag{1.10}$$

corresponds to the implicit trapezoidal rule

$$y^* = \psi_{h\,f}(y) := y + \frac{h}{2}(f(Y_1) + f(Y_2)),$$

where $Y_1 = y$, and $Y_1 \in \mathbb{R}^d$ is implicitly defined as a function of h and y by

$$Y_2 = y + \frac{h}{2}(f(Y_1) + f(Y_2)).$$

Given two RK methods specified by two different RK tableaux determining the integration maps $\psi_{h\,f}$ and $\bar{\psi}_{h\,f}$, they are *equivalent* if, when applied to an arbitrary system (1.1) with Lipschitz continuous $f : \mathcal{U} \subset \mathbb{R}^d \to \mathbb{R}^d$, for each $y_0 \in \mathcal{U}$, there exists $h_0 > 0$ such that $\psi_{h\,f}(y_0) = \bar{\psi}_{h\,f}(y_0)$ for each $h \in (-h_0, h_0)$.

For instance, it is straightforward to check that, all the RK tableaux of the one-parameter family

$$\begin{array}{|cc} 0 & 0 \\ 0 & 0 \\ \hline \lambda & 1-\lambda \end{array}$$

are equivalent to each other, and they are equivalent to the tableau (1.9). A less trivial example of equivalent RK tableaux is given by the following three-parameter family:

$$\begin{array}{|ccc} 0 & \lambda_1 & -\lambda_1 \\ 1/2 & \lambda_2 & 1/2-\lambda_2 \\ 1/2 & \lambda_3 & 1/2-\lambda_3 \\ \hline 1/2 & \lambda_4 & 1/2-\lambda_4 \end{array}$$

It can be shown that the tableaux in that family are all equivalent to each other, and that they are actually equivalent to the tableaux (1.10) corresponding to the implicit trapezoidal rule.

2 Algebraic theory of Runge-Kutta methods

2.1 The order conditions of RK methods

We now aim at deriving necessary and sufficient conditions for a RK method to be of order n when applied to arbitrary (sufficiently smooth) systems of ODEs (1.1).

In the sequel a s-stage RK method is a pair $\mu = (b, A)$ where $A = (a_{ij})_{i,j=1}^s$ is a $s \times s$ real matrix, and $b = (b_i)_{i=1}^s$ a row vector with real components, typically represented as a RK tableau (1.6). We denote the set of RK methods as $\mathcal{RK}$.

A RK method $\mu = (b, A)$ is said to be of order n ($n \geq 1$), if (1.4) holds for the integration map $\psi_{h\,f}$ (defined by (1.7)–(1.8)) associated to each smooth vector field $f : \mathcal{U} \subset \mathbb{R}^d \to \mathbb{R}^d$.

Given a s-stage RK method $\mu = (b, A)$, for each $i = 1, \ldots, s$, we denote as μ_i the RK method

$$\mu_i := \begin{array}{|ccc} a_{11} & \cdots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{s1} & \cdots & a_{ss} \\ \hline a_{i1} & \cdots & a_{is} \end{array} \,. \tag{2.1}$$

Clearly, the integration map $\psi_{h\,f}$ associated to the method $\mu = (b, A)$ is related (for each ODE system (1.1)) to the integration maps ${}^i\psi_{h\,f}$ associated to the RK methods μ_i through the relation

$$\psi_{h\,f}(y) = y + h \sum_{i=1}^{s} b_i f({}^i\psi_{h\,f}(y)). \tag{2.2}$$

Given $\mu = (b, A) \in \mathcal{RK}$, $n \geq 1$, and a smooth vector field $f : \mathcal{U} \subset \mathbb{R}^d \to \mathbb{R}^d$, in order to check whether (1.4) holds, we need to expand both $\psi_{h\,f}(y)$ and $f(\psi_{h\,f}(y))$ as series in powers of h. From (2.2), and since ${}^i\psi_{h\,f}(y) = y + \mathcal{O}(h)$, one gets that

$$\psi_{h\,f}(y) = y + h \left(\sum_{i=1}^{s} b_i \right) f(y) + \mathcal{O}(h^2).$$

Hence, $f(\psi_{h\,f}(y))$ can be expanded as

$$\begin{aligned} f(\psi_{h\,f}(y)) &= f(y) + f'(y)(\psi_{h\,f}(y) - y) + \mathcal{O}(h^2) \\ &= f(y) + h \left(\sum_{i=1}^{s} b_i \right) f'(y)(y) + \mathcal{O}(h^2). \end{aligned}$$

Since the latter expansion is valid for the integration map $\psi_{h\,f}(y)$ produced by an arbitrary RK method $\mu = (b, A)$, this also holds for the integration map ${}^i\psi_{h\,f}$ of the RK method μ_i, that is

$$f({}^i\psi_{h\,f}(y)) = f(y) + h \left(\sum_{j=1}^{s} a_{ij} \right) f'(y)(y) + \mathcal{O}(h^2).$$

By inserting that into (2.2), we in turn arrive at

$$\psi_{h\,f}(y) = y + h\Big(\sum_{i=1}^{s} b_i\Big)\, f(y) + h^2\Big(\sum_{i,j=1}^{s} b_i a_{ij}\Big)\, f'(y) f(y) + \mathcal{O}(h^3).$$

Multivariate Taylor expansion of $f(\psi_{h\,f}(y)) = f(z)$ at $z = y$ gives

$$\begin{aligned} f(\psi_{h\,f}(y)) &= f(y) + f'(y)(z-y) + \frac{1}{2} f''(y)(z-y, z-y) + \mathcal{O}(h^3) \\ &= f(y) + h\Big(\sum_{i=1}^{s} b_i\Big)\, f'(y) f(y) + h^2\Big(\sum_{i,j=1}^{s} b_i a_{ij}\Big)\, f'(y) f'(y) f(y) \\ &\quad + \frac{h^2}{2}\Big(\sum_{i=1}^{s} b_i\Big)^2 f''(y)(f(y), f(y)) + \mathcal{O}(h^3). \end{aligned}$$

where $f'(y)$ represents the Jacobian matrix of $f(y)$ with respect to y, and $f''(y)$ represent the second Fréchet derivative at y, so that $f''(y)(f(y), f(y))$ is the vector obtained by letting $f''(y)$ act on the pair $(f(y), f(y))$. That expansion is of course valid for $f({}^i\psi_{h\,f}(y))$ if the coefficients of the RK method $\mu = (b, A)$ are replaced by those of μ_i. By inserting in (2.2) the expansions of $f({}^i\psi_{h\,f}(y))$ obtained in this way, one gets an expansion of $\psi_{h\,f}(y)$ with an additional power of h. In order to write that conveniently, it will be useful to introduce some notation. From now on, $\mathbb{R}^{\mathcal{RK}}$ denotes the set of functions $u : \mathcal{RK} \to \mathbb{R}$.

Definition 2.1 *Given a function $u \in \mathbb{R}^{\mathcal{RK}}$, a new function $[u] \in \mathbb{R}^{\mathcal{RK}}$ can be defined as*

$$[u](\mu) = \sum_{i=1}^{s} b_i\, u(\mu_i).$$

We denote by e the function $e \in \mathbb{R}^{\mathcal{RK}}$ defined by $e(\mu) = 1$, $\forall \mu \in \mathcal{RK}$. Thus the function $[e] \in \mathbb{R}^{\mathcal{RK}}$ is defined by $[e](\mu) = \sum_{i=1}^{s} b_i$, for each $\mu = (b, A) \in \mathcal{RK}$.

Using that notation, we have that

$$\begin{aligned} f(\psi_{h\,f}(y)) &= f(y) + h\,[e](\mu)\, f'(y) f(y) + h^2\,[[e]](\mu)\, f'(y) f'(y) f(y) \\ &\quad + \frac{h^2}{2}\,([e](\mu))^2\, f''(y)(f(y), f(y)) + \mathcal{O}(h^3), \end{aligned}$$

and hence

$$\begin{aligned} f({}^i\psi_{h\,f}(y)) &= f(y) + h\,[e](\mu_i)\, f'(y) f(y) + h^2\,[[e]](\mu_i)\, f'(y) f'(y) f(y) \\ &\quad + \frac{h^2}{2}\,([e](\mu_i))^2\, f''(y)(f(y), f(y)) + \mathcal{O}(h^3), \end{aligned}$$

which inserted into (2.2) gives

$$\begin{aligned}
\psi_{h\,f}(y) &= y + h\,[e](\mu)\,f(y) + h^2 \sum_{i=1}^{s} b_i[e](\mu_i)\,f'(y)f(y) \\
&\quad + h^3 \sum_{i=1}^{s} b_i[[e]](\mu_i)\,f'(y)f'(y)f(y) \\
&\quad + \frac{h^3}{2} \sum_{i=1}^{s} b_i[e](\mu_i)^2\,f''(y)(f(y),f(y)) + \mathcal{O}(h^4) \\
&= y + h\,[e](\mu)\,f(y) + h^2\,[[e]](\mu)\,f'(y)f(y) \\
&\quad + h^3\,[[[e]]](\mu)\,f'(y)f'(y)f(y) \\
&\quad + \frac{h^3}{2}\,[[e]^2](\mu)\,f''(y)(f(y),f(y)) + \mathcal{O}(h^4).
\end{aligned}$$

With such expansions of $\psi_{h\,f}(y)$ and $f(\psi_{h\,f}(y))$ at hand, and using the characterization (1.4) of methods of order n, one readily obtains that a RK method $\mu = (b, A) \in \mathcal{RK}$ is of order three if

$$[e](\mu) = 1, \quad [[e]](\mu) = \frac{1}{2}\,[e](\mu), \quad [[[e]]](\mu) = \frac{1}{3}\,[[e]](\mu), \quad [[e]^2](\mu) = \frac{1}{3}\,([e](\mu))^2,$$

or equivalently,

$$[e](\mu) = 1, \quad [[e]](\mu) = \frac{1}{2}, \quad [[[e]]](\mu) = \frac{1}{6}, \quad [[e]^2](\mu) = \frac{1}{3}.$$

A systematic derivation of conditions that guarantee that a method $\mu = (b, A)$ is of order n can be obtained by generalizing the above procedure. In order to do that, we consider the following sets of functions $u \in \mathbb{R}^{\mathcal{RK}}$.

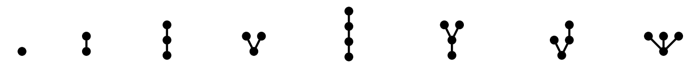
Definition 2.2 *For $n \geq 1$, consider the sets $\mathcal{T}_n \subset \mathbb{R}^{\mathcal{RK}}$ recursively defined as follows: $\mathcal{T}_1 = \{[e]\}$, and for each $n \geq 2$,*

$$\mathcal{T}_n = \Big\{[u_1 \dots u_m] \ : \ m \geq 1,\ u_i \in \mathcal{T}_{n_i} \text{ and } \sum_{i=1}^{m} n_i = n-1\Big\}$$

We also denote $\mathcal{T} = \bigcup_{n\geq 1} \mathcal{T}_n$.

Observe that for each $u \in \mathcal{T}_n$, $u(\mu)$ is a polynomial function on the entries b_i and a_{ij} in $\mu = (b, A)$. Actually, u is a function of homogeneous degree n in the following sense: Given $\mu = (b, A) \in \mathcal{RK}$, consider for each $\lambda \in \mathbb{R}$ the new method $\lambda\mu := (\lambda b, \lambda A) \in \mathcal{RK}$.[1] A function $u \in \mathbb{R}^{\mathcal{RK}}$ is of *homogeneous degree* n if for all $\lambda \in \mathbb{R}$, $u(\lambda m) = \lambda^n u(\mu)$, and we write $|u| = n$ in that case. We will

[1] If $\psi_{h\,f}$ is the integration map (defined by (1.7)–(1.8)) produced by the method μ for a smooth vector field $f : \mathcal{U} \subset \mathbb{R}^d \to \mathbb{R}^d$, then the integration map associated to the method $\lambda\mu$ is precisely $\psi_{\lambda h\,f}$, that is the map obtained from $\psi_{h\,f}$ by rescaling the time step h as λh.



$[e]$	$[[e]]$	$[[[e]]]$	$[[e]^2]$	$[[[[e]]]]$	$[[[e]^2]]$	$[[[e]][e]]$	$[[e]^3]$

Table 1: Functions $u \in \mathcal{T}$ associated to rooted trees with up to four vertices

say that a function $u \in \mathbb{R}^{\mathcal{RK}}$ is *homogeneous*, if $|u| = n$ for some non-negative integer n. Obviously, $|uv| = |u| + |v|$ if u and v are functions on the set $\mathcal{RK}$.

It was observed by Butcher that the set $\mathcal{T}$ can be identified with the set of rooted trees. Indeed, one can associate the function $[e] \in \mathcal{T}_1$ to the rooted tree with one vertex, and $[u_1 \cdots u_m] \in \mathcal{T}$ (with $u_1, \dots, u_m \in \mathcal{T}$) is associated to the rooted tree that is obtained by grafting the root of the tree corresponding to each $u_1, \dots, u_m$ to a new root. Obviously, $|u| = n$ if the function $u \in \mathcal{T}$ is associated to a rooted tree with n vertices.

The rooted trees with up to four vertices are displayed in Table 1 together with their associated functions in $\bigcup_{n=1}^{4} \mathcal{T}_n \subset \mathbb{R}^{\mathcal{RK}}$.

We can now state Butcher's original result giving necessary and sufficient order conditions for RK methods. A proof of that result is given in Subsection 2.3.

Theorem 2.3 (Butcher) *A RK method $\mu \in \mathcal{RK}$ is of order n if and only if*

$$u(\mu) = \frac{1}{u!} \quad \forall u \in \bigcup_{k=1}^{n} \mathcal{T}_k, \tag{2.3}$$

where $[e]! = 1$, and if $u = [u_1 \cdots u_m]$, with $u_1, \dots, u_m \in \mathcal{T}$,

$$u! = u_1! \cdots u_m! |u|.$$

2.2 The independence of order conditions

We will next show that, as a consequence of Theorem 2.5 below, the order conditions (2.3) given in Theorem 2.3 are all independent, in the sense that no system of algebraic equations having the same set of solutions has fewer equations than (2.3). Furthermore, Theorem 2.5 implies that the set of functions $\mathcal{T} \subset \mathbb{R}^{\mathcal{RK}}$ can actually be identified with the set of rooted trees, that is, that two functions $u, v \in \mathcal{T}$ associated to two different rooted trees are also different as functions on the set $\mathcal{RK}$ of RK methods.

We first prove an auxiliary result.

Lemma 2.4 *Given $u, v \in \mathbb{R}^{\mathcal{RK}}$, if $[u] = [v]$, then $u = v$.*

Proof This is a consequence of the following: Given $\mu = (b, A) \in \mathcal{RK}$, then the method

$$\mu' = \begin{array}{|cccc} a_{11} & \cdots & a_{1s} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{s1} & \cdots & a_{ss} & 0 \\ b_1 & \cdots & b_s & 0 \\ \hline 0 & \cdots & 0 & 1 \end{array}$$

is such that

$$\forall u \in \mathbb{R}^{\mathcal{RK}}, \quad u(\mu) = [u](\mu').$$

■

Theorem 2.5 (Butcher) *Given arbitrary $u_1, \ldots, u_k \in \mathcal{T}$ associated to k distinct rooted trees,*

$$\forall \alpha \in \mathbb{R}^k, \quad \exists \mu \in \mathcal{RK} \quad \textit{such that} \quad \begin{pmatrix} u_1(\mu) \\ \vdots \\ u_k(\mu) \end{pmatrix} = \alpha. \tag{2.4}$$

Proof We first make the following observations:

1. Given $u_1, \ldots, u_k \in \mathcal{T}$ associated to k distinct rooted trees, (2.4) implies that the functions
$$u_1^{l_1} \cdots u_k^{l_k}$$
for different k-tuples $(l_1, \ldots, l_k)$ of non-negative integers are linearly independent.

2. The set $\mathcal{RK}$ can be endowed with a vector space structure such that each $u \in \mathcal{T}$ is a linear function $u : \mathcal{RK} \to \mathbb{R}$. This is clearly achieved by defining, for given $(b', A'), (b'', A'') \in \mathcal{RK}$ and $\lambda', \lambda'' \in \mathbb{R}$, the linear combination $\lambda' (b', A') + \lambda'' (b'', A'')$ as the RK method $\mu = (b, A)$ given by
$$b = \begin{pmatrix} \lambda' b' & \lambda'' b'' \end{pmatrix}, \quad A = \begin{pmatrix} A' & 0 \\ 0 & \bar{A}'' \end{pmatrix}.$$
Since each $u_j \in \mathcal{T}$ is a linear function on the vector space $\mathcal{RK}$, (2.4) is equivalent to the linear independence of $u_1, \ldots, u_k$.

We now prove by induction on $n = \max_i |u_i|$ that the functions $u_1, \ldots, u_k \in \mathcal{T}$ associated to arbitrary k distinct rooted trees are linearly independent. The statement trivially holds for $n = 1$, as in that case $k = 1$, $u_1 = [e]$. Assume that the statement holds whenever $\max(|u_1|, \ldots, |u_k|) < n$. Hence, the functions

$$u_1^{l_1} \cdots u_k^{l_k}, \quad l_1, \ldots, l_k \geq 0,$$

are linearly independent provided that $\max(|u_1|, \ldots, |u_k|) < n$. Assume now that there exists $v_1, \ldots, v_r \in \mathcal{T}$ associated to r distinct rooted trees with

$\max(|v_1|, \dots, |v_r|) = n$ that are linearly dependent, that is, $\exists \lambda_1, \dots, \lambda_r \in \mathbb{R}$ such that

$$0 = \sum_{i=1}^{r} \lambda_i \, v_i. \tag{2.5}$$

By definition of the set $\mathcal{T}$, there exist $u_1, \dots, u_k \in \mathcal{T}$, and $l_{ij} \geq 0$ for $i = 1, \dots, r$, $j = 1, \dots, k$, such that $v_i = [u_1^{l_{i1}} \cdots u_k^{l_{ik}}]$ for each $i = 1, \dots, r$. By virtue of Lemma 2.4, we have that

$$0 = \sum_{i=1}^{r} \lambda_i \, v_i = \Big[\sum_{i=1}^{r} \lambda_i \, u_1^{l_{i1}} \cdots u_k^{l_{ik}} \Big] \quad \Longrightarrow \quad 0 = \sum_{i=1}^{r} \lambda_i \, u_1^{l_{i1}} \cdots u_k^{l_{ik}},$$

which is in contradiction with the induction hypothesis as $\max_i |u_i| < n$. ■

Remark 2.6 *Theorem 2.5 is also valid, as can be checked from its proof, with $\mathcal{RK}$ replaced by the set of explicit RK tableaux (that is, the RK tableaux (b, A) with strictly lower triangular matrices A).*

2.3 Proof of necessary and sufficient order conditions

We next give some definitions and auxiliary results that we will use to prove Theorem 2.3. Recall that the set $\mathcal{T}$ can be identified with the set of rooted trees, and we have that for each $u \in \mathcal{T}$ with $|u| > 1$, there exist a unique $m \geq 1$ and $v_1, \dots, v_m \in \mathcal{T}$, unique up to permutations, such that $u = [v_1 \cdots v_m]$.

Definition 2.7 *Given $f : \mathbb{R}^d \to \mathbb{R}^d$, we define for each $u \in \mathcal{T}$ a smooth map $F_u : \mathbb{R}^d \to \mathbb{R}^d$ (called the elementary differential of u) as follows. For $[e] = [e] \in \mathcal{T}_1$, $F_{[e]} = f$, and for $u = [u_1 \cdots u_m]$, where $u_1, \dots, u_m \in \mathcal{T}$,*

$$F_u(y) = f^{(m)}(y)(F_{u_1}(y), \dots, F_{u_m}(y)) \quad \forall y \in \mathcal{U}.$$

Definition 2.8 *We consider for $n \geq 0$ the sets $\mathcal{F}_n$ of functions on $\mathcal{RK}$ defined as follows: $\mathcal{F}_0 = \{e\}$, and for $n \geq 1$,*

$$\mathcal{F}_n = \Big\{ u_1 \dots u_m \ : \ m \geq 1, \ u_i \in \mathcal{T}_{n_i} \ \textit{and} \ \sum_{i=1}^{m} n_i = n \Big\}.$$

We also denote $\mathcal{F} = \bigcup_{n \geq 0} \mathcal{F}_n$.

Recall that $e \in \mathbb{R}^{\mathcal{RK}}$ is defined by $e(\mu) = 1$. From the proof of Theorem 2.5, one gets that each $u \in \mathcal{F}$ can be uniquely factored in the form $u = u_1^{l_1} \cdots u_k^{l_k}$, where $u_1, \dots, u_k \in \mathcal{T}$ are distinct functions and $l_1, \dots, l_k \geq 1$.

It is straightforward to check that each $u \in \mathcal{F}_n$ has homogeneous degree n. Comparing the definition of the sets $\mathcal{T}_n$ and $\mathcal{F}_{n-1}$, one sees that, for each $n \geq 1$, $\mathcal{T}_n = \{[u] \ : \ u \in \mathcal{F}_{n-1}\}$. Clearly,

$$\mathcal{F} = \{e\} \cup \Big\{ u_1 \dots u_m \ : \ m \geq 1, \ u_i \in \mathcal{T} \Big\}.$$

$\lvert u\rvert$	1	2	3	3	4	4	4	4
u								
F_u	f	$f'f$	$f'f'f$	$f''(f,f)$	$f'f'f'f$	$f'f''(f,f)$	$f''(f'f,f)$	$f'''(f,f,f)$
$u!$	1	2	6	3	24	12	8	4
$\sigma(u)$	1	1	1	2	1	2	1	6

Table 2: Elementary differentials F_u and the values of $u!$ and $\sigma(u)$ for rooted trees u with up to four vertices

The set of functions $\mathcal{F} \subset \mathbb{R}^{\mathcal{RK}}$ can be identified with the set of forests of rooted trees. A forest is a collection of (possibly repeated) rooted trees.

Definition 2.9 *We recursively define a positive integer for each $u \in \mathcal{F}$, $\sigma(u)$ (the* symmetry number *of u) as follows: $\sigma(e) = 1$, $\sigma([u]) = \sigma(u)$ if $u \in \mathcal{F}$, and if $u_1, \ldots, u_m \in \mathcal{T}$ are distinct functions and $l_1, \ldots, l_m \geq 1$,*

$$\sigma(u_1^{l_1} \cdots u_m^{l_m}) = l_1! \cdots l_m! \sigma(u_1)^{l_1} \cdots \sigma(u_m)^{l_m}.$$

The elementary differentials F_u, and the values of $u!$ and $\sigma(u)$ associated to rooted trees with four or less vertices are displayed in Table 2.

Proposition 2.10 *Consider an ODE system (1.1) and the integration map $\psi_{h\,f}$ associated to an arbitrary RK method $\mu \in \mathcal{RK}$. Then it holds that, for each $y \in \mathcal{U}$, $\psi_{h\,f}(y)$ can be expanded as follows,*

$$\begin{aligned} \psi_{h\,f}(y) &= y + \sum_{u \in \mathcal{T}} \frac{h^{|u|}}{\sigma(u)} u(\mu)\, F_u(y), \\ &= y + \sum_{n=1}^{\infty} h^n \sum_{u \in \mathcal{T}_n} \frac{1}{\sigma(u)} u(\mu)\, F_u(y). \end{aligned} \tag{2.6}$$

In addition, $f(\psi_{h\,f}(y))$ can be expanded as

$$f(\psi_{h\,f}(y)) = f(y) + \sum_{v \in \mathcal{F}} \frac{h^{|v|}}{\sigma(v)} v(\mu)\, F_{[v]}(y). \tag{2.7}$$

Proof We will first show that (2.6) formally implies (2.7). In order to do that, we resort to the multivariate Taylor expansion of $f(z)$ at $z = y$,

$$f(z) = f(y) + \sum_{m \geq 1} \frac{1}{m!} f^{(m)}(y)(\overbrace{z - y, \ldots, z - y}^{m \text{ times}}). \tag{2.8}$$

Here, $f^{(m)}(y)$ is the mth Fréchet derivative of f at y. We replace z in (2.8) by the expansion (2.6), and taking the multilinearity of $f^{(m)}(y)$ into account, we obtain that $f(z) - f(y)$ can be expanded as

$$\sum_{m\geq 1} \sum_{u_1,\ldots,u_m\in\mathcal{T}} \frac{h^{|u_1|+\cdots+|u_m|}}{m!} \frac{u_1(\mu)\cdots u_m(\mu)}{\sigma(u_1)\cdots\sigma(u_m)} f^{(m)}(y)(F_{u_1}(y),\ldots,F_{u_m}(y))$$

and by the definition of elementary differentials F_u, we get that $f(\psi_{h\,f}(y))$ admits the expansion

$$f(y) + \sum_{m\geq 1} \sum_{u_1,\ldots,u_m\in\mathcal{T}} \frac{h^{|u_1\cdots u_m|}}{m!} \frac{(u_1\cdots u_m)(\mu)}{\sigma(u_1)\cdots\sigma(u_m)} F_{[u_1\cdots u_m]}(y). \tag{2.9}$$

Collecting the repeated terms, a term per $v = u_1\cdots u_m \in \mathcal{F}$ is obtained. If $v_1,\ldots,v_k$ are the distinct rooted trees in $\{u_1,\ldots,u_m\}$, and the number of rooted trees u_j that coincide with v_i is denoted by l_i, so that $u_1\cdots u_m = v_1^{l_1}\cdots v_k^{l_k}$, then the term corresponding to $v = u_1\cdots u_m$ appears $m!/(l_1!\cdots l_k!)$ times in (2.9). This shows, by virtue of the definition of $\sigma(v)$ that $f(\psi_{h\,f}(y))$ can be expanded as (2.7).
Thus, if (2.6) holds up to terms of degree n in h for arbitrary RK methods $\mu \in \mathcal{RK}$, then (2.7) also holds up to terms of degree n in h. Then, (1.7) implies that (2.6) holds up to terms of degree $n+1$ in h. ∎

For each $n \geq 1$, let d_n be the cardinality of the set $\mathcal{T}^n := \cup_{k=1}^n \mathcal{T}_k$. Consider a total ordering $u_1 < u_2 < u_3 < \cdots$ of the set $\mathcal{T}$ such that, for each $n \geq 0$, $\mathcal{T}_{n+1} = \{u_{d_n+1},\ldots,u_{d_{n+1}}\}$. The following result can be proven by induction on n.

Lemma 2.11 *Given $n \geq 1$, consider the polynomial vector field $f : \mathbb{R}^{d_n} \to \mathbb{R}^{d_n}$ defined as follows: For $y = (y^1,\ldots,y^{d_n})^T$, $f(y) = (f^1(y),\ldots,f^{d_n}(y))^T$ where $f^1(y) = 1$, and for $i \geq 2$, if $u_i = [u_{j_1}^{l_1}\cdots u_{j_m}^{l_m}]$ where $u_{j_1},\ldots,u_{j_m} \in \mathcal{T}$ are distinct and $l_1,\ldots,l_m \geq 1$, then*

$$f^i(y) = \frac{1}{(l_1+\cdots+l_m)!}(y^{j_1})^{l_1}\cdots(y^{j_m})^{l_m}.$$

The elementary differentials F_{u_i}, $u_i \in \mathcal{T}^n$, corresponding to f evaluated at the origin $0 \in \mathbb{R}^{d_n}$ satisfy

$$F_{u_i}^j(0) = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases}$$

Proof of Theorem 2.3 Application of Proposition 2.10 and the characterization (1.4) of methods of order n lead to the following sufficient conditions for order n,

$$[v](\mu) = \frac{1}{k} v(\mu) \quad \forall v \in \mathcal{F}_{k-1}, \quad 1 \leq k \leq n. \tag{2.10}$$

Lemma 2.11 implies that condition (2.10) is also necessary. Since by definition, $\mathcal{T}_k = \{[v] \ : \ \mathcal{F}_{k-1}\}$, the required result follows by observing that, by virtue of the recursive definition of $u!$, (2.10) is equivalent to (2.3). ∎

2.4 Composition of RK methods, rooted trees, and forests

We next define a composition in $\mathcal{RK}$ that is compatible (for each ODE (1.1)) with the composition of integration maps $\psi_{h\,f}$ associated to RK methods.

Definition 2.12 *Given a s-stage RK method $\mu = (b, A) \in \mathcal{RK}$ and a s-stage RK method $\mu' = (b', A') \in \mathcal{RK}$, the composition $\mu\,\mu' \in \mathcal{RK}$ is the $(s+s')$-stage RK method given by the Butcher tableau*

$$\begin{array}{|cccccc} a_{11} & \cdots & a_{1s} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{s1} & \cdots & a_{ss} & 0 & \cdots & 0 \\ b_1 & \cdots & b_s & a'_{11} & \cdots & a'_{1s'} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ b_1 & \cdots & b_s & a'_{s'1} & \cdots & a'_{s's'} \\ \hline b_1 & \cdots & b_s & b'_1 & \cdots & b'_{s'} \end{array} .$$

Clearly, if $\psi_{h\,f}$ and $\psi'_{h\,f}$ are the integration maps associated the RK methods μ and μ' respectively, then the integration map associated to $\mu\,\mu'$ is the composition $\psi'_{h\,f} \circ \psi_{h\,f}$.

Proposition 2.13 *Given an arbitrary function $u \in \mathbb{R}^{\mathcal{RK}}$, for all $\mu, \mu' \in \mathcal{RK}$,*

$$[u](\mu\,\mu') = [u](\mu) + [u_\mu](\mu'), \tag{2.11}$$

where $u_\mu \in \mathbb{R}^{\mathcal{RK}}$ is defined as $u_\mu(\mu') = u(\mu\,\mu')$ for all $\mu' \in \mathcal{RK}$.

Proof We first note that, from Definition 2.12,

$$(\mu\,\mu')_i = \begin{cases} \mu_i & \text{if} \quad 1 \le i \le s, \\ \mu'_{i-s}\,\mu & \text{if} \quad s+1 \le i \le s+s'. \end{cases}$$

Given $u \in \mathbb{R}^{\mathcal{RK}}$, by applying Definition 2.1 to $[u] \in \mathbb{R}^{\mathcal{RK}}$ we obtain

$$\begin{aligned} [u](\mu\,\mu') &= \sum_{i=1}^{s} b_i u((\mu\,\mu')_i) + \sum_{i=1}^{s'} b'_i u((\mu\,\mu')_{s+i}) \\ &= \sum_{i=1}^{s} b_i u(\mu_i) + \sum_{i=1}^{s'} b'_i u(\mu\,\mu'_i) \\ &= \sum_{i=1}^{s} b_i u(\mu_i) + \sum_{i=1}^{s'} b'_i u_\mu(\mu'_i), \end{aligned}$$

and (2.11) follows from applying Definition 2.1 to $[u](\mu)$ and $[u_\mu](\mu')$. ∎

The identity (2.11) can be equivalently written as

$$[u]_\mu = [u](\mu)\,e + [u_\mu] \tag{2.12}$$

which allows recursively writting formulae for $u(\mu\mu')$ for $u \in \mathcal{F}$, starting from $(e)_\mu = e$. For instance,

$$
\begin{aligned}
[e]_\mu &= [e](\mu)\, e + [e], \\
[[e]]_\mu &= [[e]](\mu)\, e + [[e](\mu)\, e + [e]] \;=\; [[e]](\mu)\, e + [e](\mu)\, [e] + [[e]], \\
([e]^2)_\mu &= ([e](\mu)\, e + [e])^2 \;=\; [e](\mu)^2\, e + 2[e](\mu)\, [e] + [e]^2, \\
[[e]^2]_\mu &= [[e]^2](\mu)\, e + [[e](\mu)^2\, e + 2[e](\mu)\, [e] + [e]^2] \\
&= [[e]^2](\mu)\, e + [e](\mu)^2\, [e] + 2[e](\mu)\, [[e]] + [[e]^2], \\
[[[e]^2]]_\mu &= [[[e]^2]](\mu)\, e + [[[e]^2]\, e + [e](\mu)^2\, [e] + 2[e](\mu)\, [[e]] + [[e]^2]] \\
&= [[[e]^2]](\mu)\, e + [[e]^2](\mu)\, [e] + [e](\mu)^2\, [[e]] + 2[e](\mu)\, [[[e]]] + [[[e]^2]].
\end{aligned}
$$

or equivalently, rewriting $u_\mu(\mu')$ back as $u(\mu\,\mu')$ and representing each $u \in \mathcal{T}$ by the corresponding rooted tree,

$$
\begin{aligned}
\bullet(\mu\,\mu') &= \bullet(\mu) + \bullet(\mu'), \\
[[e]](\mu\,\mu') &= [[e]](\mu) + \bullet(\mu)\,\bullet(\mu') + [[e]](\mu'), \\
\Big(\bullet(\mu\,\mu')\Big)^2 &= \Big(\bullet(\mu)\Big)^2 + 2\bullet(\mu)\,\bullet(\mu') + \Big(\bullet(\mu')\Big)^2, \\
[[e]^2](\mu\,\mu') &= [[e]^2](\mu) + \Big(\bullet(\mu)\Big)^2 \bullet(\mu') + 2\bullet(\mu)\,[[e]](\mu') + [[e]^2](\mu'), \\
[[[e]^2]](\mu\,\mu') &= [[[e]^2]](\mu) + [[e]^2](\mu)\,\bullet(\mu') + \Big(\bullet(\mu)\Big)^2 [[e]](\mu') \\
&\quad + 2\bullet(\mu)\,[[[e]]](\mu') + [[[e]^2]](\mu').
\end{aligned} \tag{2.13}
$$

The formulae in (2.13) for $u(\mu\,\mu')$ can be interpreted in terms of the rooted tree associated to u. For instance, $u(\mu\,\mu')$ for the rooted tree $u = [[[e]^2]]$ is a sum of terms of the form $v(\mu)w(\mu')$ where $(v, w) \in \mathcal{F} \times \mathcal{T}$ is obtained from pruning the original tree (where w is the rooted tree that remains after pruning, and v corresponds to the pieces that are removed). In that example, the collection of such (v, w) is

$$([[[e]^2]], e), ([[e]^2], \bullet), (\bullet\bullet, [[e]]), (\bullet, [[[e]]]), (\bullet, [[[e]]]), (e, [[[e]^2]]).$$

We will next make this precise in the general case.

The diagrams representing rooted trees in Table 1 can be identified with partially ordered sets U of points in the plane (the vertices of U) having only one minimal vertex (the root of U) and satisfying the following:

$$x, y, z \in U \quad x < z, \quad y < z \quad \Longrightarrow \quad x < y \quad \text{or} \quad y < x. \tag{2.14}$$

Actually, a rooted tree can be defined as an isomorphism class[2] of finite partially ordered sets satisfying (2.14) and having only one minimal vertex (the

[2]An isomorphism of partially ordered sets is a bijection of the sets that preserve their partial ordering

root). In turn, a forest of rooted trees can be defined as an isomorphism class of finite partially ordered sets satisfying (2.14). Given two disjoint partially ordered sets V and W representing two forests v and w respectively, its (disjoint) union $U = V + W$ also represents a forest, which we denote by $u = vw$. Clearly, any partially ordered set U satisfying (2.14) and having $m > 1$ minimal vertices, can be uniquely decomposed as the disjoint union of m partially ordered sets $U_1, \ldots, U_m$ representing rooted rees (i.e., satisfying (2.14) and having only one minimal vertex), and in that case, U represents the forest $u = u_1 \cdots u_m$.

Given a rooted tree v, one obviously obtains a forest u by removing the root of v. We will write $v = [u]$ in that case (thus mimicking the notation we have used so far for the associated functions in $\mathbb{R}^{\mathcal{RK}}$). Any rooted tree can thus be represented as $[u]$, where u is a uniquely determined forest of rooted trees. Later on, we will make use of an additional operation on rooted trees and forests, the so-called (left) Butcher product: Given a rooted tree $v = [u]$ (u a forest) and a forest w, we denote by $w \bullet v = [uw]$, that is, $w \bullet v$ is the rooted tree of degree $|v| + |w|$ obtained by grafting each of the labeled rooted trees in the forest w to the root of v.

Given a partially ordered set U, we will say that a pair $(x, y) \in U \times U$ is comparable in U if either $x < y$ or $y < x$. Given partially ordered subsets $V_1, \ldots, V_m$ of U, we write $(V_1 \succ \cdots \succ V_m) \subset U$ if the following three conditions are satisfied:

- each V_i is a partially ordered subset of U,
- as a set, U is the disjoint union of $V_1, \ldots, V_m$,
- if $(x, y) \in V_i \times V_{i+1}$ is comparable in U, then $x > y$.

Theorem 2.14 (Butcher) *Given $u \in \mathcal{F}$, let U be a partially ordered set representing the forest associated to u. Then $\forall (\mu, \mu') \in \mathcal{RK} \times \mathcal{RK}$,*

$$u(\mu\, \mu') = \sum_{(V \succ W) \subset U} v(\mu)\, w(\mu') \tag{2.15}$$

where the partially ordered sets V and W represent the forests associated to the functions v and w respectively.

Proof We first observe that, using the notation in Proposition 2.13, (2.15) can equivalently be written as

$$u_\mu = \sum_{(V \succ W) \subset U} v(\mu)\, w. \tag{2.16}$$

We will prove by induction on n that (2.16) holds true for all $u \in \mathcal{F}$ with $|u| \leq n$. This trivially holds for $n = 0$, as $\mathcal{F}_0 = \{e\}$, and $e_\mu = e$, while e corresponds to the empty forest, represented by the empty partially ordered set $\emptyset$, which admits the unique decomposition $(\emptyset \succ \emptyset) \subset \emptyset$.

Assume now that for all $u \in \mathcal{F}$ with $|u| \leq n-1$, (2.16) holds. Then (2.12) implies that

$$[u]_\mu = [u](\mu)\, e + \sum_{(V \succ W) \subset U} v(\mu)\, [w], \tag{2.17}$$

where we have used that, by virtue of Definition 2.1, $[\cdot] : \mathbb{R}^{\mathcal{RK}} \rightarrow \mathbb{R}^{\mathcal{RK}}$ is linear. If U' is the partially ordered set representing $[u]$, and $x \in U'$ is the root of U', then the partially ordered set $U = U' \backslash \{x\}$ represents u, and the pairs of partially ordered sets $(V', W') \neq (U, \emptyset)$ such that $(V' \succ W') \subset U'$ are in one-to-one correspondence with the pairs (V, W) such that $(V \succ W) \subset U$, with $V = V'$ and $W = W' \backslash \{x\}$. This together with (2.17) implies that

$$[u]_\mu = \sum_{(V \succ W') \subset U'} v(\mu)\, w'.$$

We thus have that (2.16) is true for all $u \in \mathcal{T}$ with $|u| \leq n$. Given $u \in \mathcal{F} \backslash \mathcal{T}$ with $|u| = n$, consider $u', u'' \in \mathcal{F} \backslash \{e\}$ such that $u = u'u''$. If U is a partially ordered set representing u, then U must be the disjoint union of two partially ordered sets U' and U'' representing u' and u'' respectively. Then, by application of the induction hypothesis, we have that

$$u_\mu = \sum_{\substack{(V' \succ W') \subset U' \\ (V'' \succ W'') \subset U''}} v'(\mu)\, v''(\mu)\, w'\, w'',$$

which finally leads to (2.16) due to the following property of partially ordered sets satisfying (2.14): Given two disjoint partially ordered sets U' and U'' and $U = U' + U''$, if $(V \succ W) \subset U$, then there exist unique partially ordered sets V', V'', W', W'' satisfying that $V = V' + V''$, $W = W' + W''$, $(V' \succ W') \subset U'$, $(V'' \succ W'') \subset U''$. ■

2.5 The Butcher group

From now on, $\mathcal{T}$ and $\mathcal{F}$ will denote the set of rooted trees and the set of forests of rooted trees respectively.

Following Butcher [2], we will define a product $*$ on the set $\mathbb{R}^{\mathcal{F}}$ of functions

$$\begin{array}{rccc} \alpha : & \mathcal{F} & \longrightarrow & \mathbb{R} \\ & u & \mapsto & \alpha_u \end{array}$$

that will endow $\mathbb{R}^{\mathcal{F}}$ with a unital (non-commutative) associative algebra structure closely related to the composition of RK methods.

Given $\mu, \mu' \in \mathcal{RK}$, consider $\alpha, \beta, \gamma \in \mathbb{R}^{\mathcal{F}}$ such that, $\forall u \in \mathcal{F}$,

$$\alpha_u = u(\mu), \quad \beta_u = u(\mu'), \quad \gamma_u = u(\mu\, \mu').$$

With that notation, (2.15) can be rewritten as $\gamma = \alpha * \beta$, where the product $*$ is defined next.

Definition 2.15 *Given $\alpha, \beta \in \mathbb{R}^{\mathcal{F}}$, $\alpha * \beta \in \mathbb{R}^{\mathcal{F}}$ is defined as follows: If U is a partially ordered set representing a forest $u \in \mathcal{F}$,*

$$(\alpha * \beta)_u = \sum_{(V \succ W) \subset U} \alpha_V \, \beta_W, \tag{2.18}$$

where $\alpha_V = \alpha_v$ and $\beta_W = \beta_w$ if V and W represent the forests v and w respectively.

The product $*$ is well defined in $\mathbb{R}^{\mathcal{F}}$ (that is, $(\alpha * \beta)_u$ does not depend on the partially ordered set U chosen as a representative of u). Indeed, if U and U' are partially ordered sets representing the same forest $u \in \mathcal{F}$, then $(V \succ W) \subset U$ implies that $(V' \succ W') \subset U'$, where the isomorphism from U to U' sends V and W to V' and W' respectively.

The product $*$ endows the vector space $\mathbb{R}^{\mathcal{F}}$ with a unital associative algebra structure with unit $\mathbb{1} \in \mathbb{R}^{\mathcal{F}}$ defined as $\mathbb{1}(e) = 1$ and $\mathbb{1}_u = 0$ if $u \in \mathcal{F} \backslash \{e\}$. Observe that the associativity of $*$ is immediate from his definition, since

$$(\alpha * (\beta * \gamma))_u = ((\alpha * \beta) * \gamma)_u = \sum_{(V_1 \succ V_2 \succ V_3) \subset U} \alpha_{v_1} \beta_{v_2} \gamma_{v_3}.$$

An important property of the product in Definition 2.15 is that, for each $u \in \mathcal{F}$, $(\alpha * \beta - \beta_e \alpha - \alpha_e \beta)_u$ is a $\mathbb{Z}$-linear combination of terms of the form $\alpha_v \beta_w$ with $\max(|v|, |w|) < |u|$, which allow proving results by induction on the number of vertices $|u|$ of the forest u.

For instance, one can easily prove that, given $\alpha \in \mathbb{R}^{\mathcal{F}}$, the equation $\alpha * \beta = \mathbb{1}$ can be uniquely solved for $\beta \in \mathbb{R}^{\mathcal{F}}$ provided that $\alpha_e \neq 0$. This proves that the subset $\{\alpha \in \mathbb{R}^{\mathcal{F}} \ : \ \alpha_e \neq 0\} \subset \mathbb{R}^{\mathcal{F}}$ is a group under the product $*$. Such a group contains an important subgroup, the *Butcher group*:

Proposition 2.16 (Butcher) *The subset of $\mathbb{R}^{\mathcal{F}}$*

$$\mathcal{G} = \{\alpha \in \mathbb{R}^{\mathcal{F}} \ : \ \alpha_e = 1, \ \forall u, v \in \mathcal{F} \ \alpha_{uv} = \alpha_u \alpha_v\}, \tag{2.19}$$

is a group under the product $*$ with identity element $\mathbb{1}$.

Proof It is sufficient to prove that $\mathcal{G}$ is a subgroup of $\{\alpha \in \mathbb{R}^{\mathcal{F}} \ : \ \alpha_e \neq 0\}$. Indeed, one can prove that $\alpha * \beta \in \mathcal{G}$ if $\alpha, \beta \in \mathcal{G}$ by following the argument used at the end of the proof of Theorem 2.14. ■

Obviously, each element $\alpha \in \mathcal{G}$ is uniquely determined by its values α_u for $u \in \mathcal{T}$, so that, as a set, $\mathcal{G}$ can be identified with the set of functions $\mathbb{R}^{\mathcal{T}}$.

We end this subsection by stating two fundamental properties of the product (2.18):

- P1 The restriction to $\mathcal{T}$ of $\alpha * \beta - \beta_e \alpha$ depends linearly on the restriction to $\mathcal{T}$ of β. This is a consequence, according to Definition 2.15, of the fact that, if U is a partially ordered set representing a rooted tree u, and $(V \succ W) \subset U$ with $W \neq \emptyset$, then W necessarily represents a rooted tree.

P2 Given $\beta \in \mathbb{R}^{\mathcal{F}}$ with $\beta_e = 0$ (and in particular, $\beta = \alpha - \mathbb{1}$ with $\alpha \in \mathcal{G}$), $(\beta^{*n})_u = 0$ provided that $u \in \mathcal{F}$ with $|u| < n$. Indeed, in that case,

$$(\beta^{*n})_u = \sum_{(V_1 \succ \cdots \succ V_n) \subset U} \beta_{V_1} \cdots \beta_{V_n}$$

where U is a partially ordered set representing a rooted tree u, but since $|u| < n$, some of the V_j must necessarily be the empty partially ordered set (thus representing the empty forest e), which implies that $(\beta^{*n})_u = 0$.

2.6 Equivalence classes of RK methods

Recall that two RK methods $\mu = (b, A)$, $\mu' = (b', A')$ may be equivalent (Subsection 1.2) in the sense of producing for each system (1.1) exactly the same numerical solution for sufficiently small step-sizes (provided that f is Lipschitz continuous).

Proposition 2.10 implies that, if $\mu, \mu' \in \mathcal{RK}$ are equivalent, then $u(\mu) = u(\mu')$ for all $u \in \mathcal{T}$. Actually, as shown by Butcher [2] (see also [3]), the reverse also holds true.

Given a RK method $\mu \in \mathcal{RK}$, let us as consider

$$\begin{array}{rccc} \widehat{\mu}: & \mathcal{F} & \longrightarrow & \mathbb{R} \\ & u & \mapsto & u(\mu). \end{array} \tag{2.20}$$

Then, the set of equivalence classes of RK methods can be identified with the set

$$\mathcal{G}_{RK} = \{\widehat{\mu} \ : \mu \in \mathcal{RK}\} \subset \mathbb{R}^{\mathcal{F}}. \tag{2.21}$$

It is clear that the set $\mathcal{G}_{RK} \subset \mathbb{R}^{\mathcal{F}}$ of equivalence classes of RK methods is contained in $\mathcal{G}$. Actually, as shown in [2], $\mathcal{G}_{RK}$ is a subgroup of $\mathcal{G}$.

Moreover, Theorem 2.5 implies that $\mathcal{G}_{RK}$ is dense in $\mathcal{G}$, with the structure of a topological group determined by the neighbourhood basis $\{\mathcal{U}_n\}_{n\geq 1}$ at $\mathbb{1} \in \mathcal{G}$ given by $\mathcal{U}_n = \{\alpha \in \mathcal{G} \ : \alpha_u = 0 \text{ if } u \in \mathcal{T}_k \text{ with } k \leq n\}$. (Furthermore, $\mathcal{U}_1 \subset \mathcal{U}_2 \subset \cdots$ are normal subgroups of $\mathcal{G}$.)

2.7 Bibliographical comments

Essentially all the results in Section 2 are due to Butcher [2]. Butcher's original proof of Theorem 2.14 makes use of a different (although equivalent) recursion based on the Butcher product $w \bullet v$ of two rooted trees v and w defined in Subsection 2.4 (in Butcher's original work, $v \cdot w$ denotes what we write as $w \bullet v$). The recursive formula (2.12) (written with a different notation) can be found in [7, 4] as an equivalent formulation of Butcher's original recursion. In Butcher's original definition of the product (2.18), subsets of a particular large partially ordered set is used instead of arbitrary partially ordered sets.

3 B-series and related formal expansions

In [19], it is observed that there are integration methods other than RK methods (for instance, multiderivative Runge-Kutta methods) that admit expansions that generalize those in Proposition 2.10: the so-called B-series (i.e., Butcher series). Working with B-series allows working directly with equivalence classes of Runge-Kutta methods and other one-step methods. The B-series expansion of one step of such a method is a series in powers of h that is convergent for sufficiently small h in the case of real analytic vector fields f. In addition, B-series that do not necessarily correspond to the expansion of actual integration methods (and may not converge even in the real analytic case) have proven to be very useful in the analysis of numerical methods for ODEs.

3.1 B-series

Following [19], we define for each $\alpha \in \mathbb{R}^{\mathcal{T}\cup\{e\}}$ a formal series as follows.

Definition 3.1 (Hairer and Wanner) *Given a vector field f of an ODE (1.1), the B-series associated to $\alpha \in \mathbb{R}^{\mathcal{T}\cup\{e\}}$ is defined as*

$$B_{hf}(\alpha, y) = \alpha_e\, y + \sum_{u\in\mathcal{T}} \frac{h^{|u|}}{\sigma(u)} \alpha_u\, F_u(y). \tag{3.1}$$

We will drop the subindex hf from $B_{hf}(\alpha, y)$ when there is no ambiguity.

With that notation, Proposition 2.10 can be restated as follows: the integration map $\psi_{hf}(y)$ of an arbitrary RK method $\mu \in \mathcal{RK}$ can be expanded as a B-series $B_{hf}(\alpha, y)$, where for each $u \in \mathcal{T} \cup \{e\}$, $\alpha_u = u(\mu)$ (and in particular, $\alpha_e = 1$), and $f(\psi_{hf}(y))$ can be expanded as a B-series $B_{hf}(\beta, y)$, where $\beta_e = 0$, and for each $u \in \mathcal{T}$, with $u = [v]$, $v \in \mathcal{T}$, $\beta_u = v(\mu)$.

Given two RK methods $\mu, \mu' \in \mathcal{RK}$ and their composition $\mu\mu' \in \mathcal{RK}$, consider their corresponding B-series $B(\alpha, y)$, $B(\beta, y)$, and $B(\gamma, y)$, that is, for each $u \in \mathcal{T} \cup \{e\}$, $\alpha_u = u(\mu)$, $\beta_u = u(\mu')$, $\gamma_u = u(\mu\mu')$ (in particular, $\alpha_e = \beta_e = \gamma_e = 1$), then we must formally have that

$$B(\beta, B(\alpha, y)) = B(\gamma, y).$$

Theorem 2.14 states that, using the notation introduced in Definition 2.15, for each $u \in \mathcal{T} \cup \{e\}$, $\gamma_u = (\alpha * \beta)_u$, where for each forest $v \in \mathcal{F}$, $\alpha_v = v(\mu)$, that is

$$\alpha_{u_1\cdots u_m} = \alpha_{u_1} \cdots \alpha_{u_m}, \quad \text{if} \quad v = u_1 \cdots u_m \in \mathcal{F}. \tag{3.2}$$

(Recall that, according to Property P2 in Subsection 2.5, the values α_v and β_w for $(v, w) \in \mathcal{F} \times (\mathcal{T} \cup \{e\})$ are required to determine all the values $(\alpha * \beta)_u$, $u \in \mathcal{T} \cup \{e\}$.)

Furthermore, Theorem 2.5 implies that, for arbitrary $\alpha, \beta \in \mathbb{R}^{\mathcal{T}\cup\{e\}}$ with $\alpha_e = \beta_e = 1$,

$$B(\beta, B(\alpha, y)) = B(\alpha * \beta, y) \tag{3.3}$$

where $\alpha \in \mathbb{R}^{\mathcal{T}\cup\{e\}}$ is extended to $\alpha \in \mathcal{G} \subset \mathbb{R}^{\mathcal{F}}$ by (3.2). Observe that (3.2) defines a bijection from the set $\{\alpha \in \mathbb{R}^{\mathcal{T}\cup\{e\}} \ : \alpha_e = 1\}$ to $\mathcal{G}$.

Observe that $\mathbb{1} \in \mathcal{G}$ corresponds, as expected, to the B-series representing the identity map, that is, $B(e, y) \equiv y$. Formula (3.3) shows that a B-series $B_{hf}(\alpha, y)$ (for a given vector field f) defines a homomorphism from the group $\mathcal{G}$ to the group of near-to-identity formal maps $\psi_h(y) = y + hg_1(y) + h^2 g_2(y) + \cdots$ in $\mathbb{R}^d$ (with smooth maps $g_n : \mathbb{R}^d \to \mathbb{R}^d$).

The composition formula (3.3) can also be extended to the case $\beta_e \neq 1$ leading to the following result. A direct proof (that does not make use of the results in Section 2) was given in the original paper [19]. An alternative direct proof of a more general results (Theorem 3.9) will be given in Subsection 3.3.

Theorem 3.2 (Hairer and Wanner) *Given two B-series $B(\alpha, y)$ and $B(\beta, y)$, where $\alpha, \beta \in \mathbb{R}^{\mathcal{T}\cup\{e\}}$ with $\alpha_e = 1$, (3.3) formally holds, where $\alpha * \beta \in \mathbb{R}^{\mathcal{F}}$ is given by (2.18) and (3.2).*

3.2 Backward error analysis, the exponential, and the logarithm

The problem of standard backward error analysis of numerical integration methods for ODEs can be stated as follows: Given $\alpha \in \mathcal{G}$ representing an integration method that, when applied to the ODE (1.1) with solution $y(t)$, provides the approximations $y_n \approx y(hn)$ at $t = nh$ in a step-by-step manner as

$$y_{n+1} = B_{hf}(\alpha, y_n), \quad n = 0, 1, 2, \ldots, \tag{3.4}$$

determine a *modified ODE* of the form

$$\frac{d\tilde{y}}{dt} = \tilde{f}(\tilde{y}; h), \quad \tilde{y}(0) = y_0, \tag{3.5}$$

(which is expected to be a perturbation of the ODE (1.1) parameterized by the discretization parameter h), whose solution $\tilde{y}(t)$ satisfies that $\tilde{y}(nh) = y_n$ for all n, or equivalently,

$$\tilde{y}(nh) = B_{hf}(\alpha^{*n}, y_0), \quad n = 0, 1, 2, \ldots \tag{3.6}$$

Let us assume that such a modified equation exists, and admits an expansion in powers of h of the form

$$\tilde{f}(\tilde{y}; h) = \tilde{f}_1(\tilde{y}) + h\tilde{f}_1(\tilde{y}) + h^2 \tilde{f}_2(\tilde{y}) + \cdots$$

In that case, application of polynomial interpolation with nodes $t = nh$, $n = 0, \ldots, N$ (for an arbitrarily high positive integer N) gives an interpolant $P_N(t)$ of the solution $\tilde{y}(t)$ of the modified equation (3.5) having a B-series expansion $P_N(t) = B_{hf}(\gamma^{[N]}(t/h), y_0)$, where $\forall u \in \mathcal{T}$, $\gamma^{[N]}(\tau)_u$ is the unique polynomial of degree N in τ satisfying that

$$\gamma^{[N]}(n)_u = (\alpha^{*n})_u \quad \text{for} \quad n = 0, 1, \ldots, N. \tag{3.7}$$

Application of Newton's forward finite difference formula for the interpolating polynomial gives

$$\gamma^{[N]}(\tau) = \mathbb{1} + \sum_{n\geq 1}^{N} \frac{\tau(\tau-1)\cdots(\tau-n+1)}{n!}(\alpha-\mathbb{1})^{*n}.$$

Standard error analysis of polynomial interpolation shows that the series in powers of h of $\tilde{y}(t)$ and $P_N(t)$ coincide up to terms of degree N in h. By considering arbitrarily high N, one gets that

$$\tilde{y}(t) = B_{hf}(\gamma(t/h), y_0), \tag{3.8}$$

where

$$\gamma(\tau) = \mathbb{1} + \sum_{n\geq 1} \frac{\tau(\tau-1)\cdots(\tau-n+1)}{n!}(\alpha-\mathbb{1})^{*n}. \tag{3.9}$$

Observe that $\gamma(\tau)$ is well defined as a function in $\mathbb{R}^{\mathcal{F}}$, because $\alpha_e = \mathbb{1}$ implies that $((\alpha_e-\mathbb{1})^{*n})_u = 0$ whenever $|u| < n$ (Property P2 in Subsection 2.5).

If $\tilde{y}(t)$ given by (3.8)–(3.9) is the solution of an autonomous ODE (3.5), then the right-hand side of (3.5) can be recovered from the expansion (3.8) of its solution $\tilde{y}(t)$ as

$$\tilde{f}(y_0;h) = \frac{d}{dt}\tilde{y}(t)\Big|_{t=0} = \frac{d}{dt}B_{hf}(\gamma(t/h), y_0)\Big|_{t=0},$$

that is,

$$\begin{aligned} \tilde{f}(y;h) &= h^{-1}B_{hf}(\beta, y) \\ &= \beta_{\bullet}\, f(y) + h\,\beta_{[\bullet]}\, f'(y)f(y) + \cdots \end{aligned} \tag{3.10}$$

where the coefficients β_u for $u \in \mathcal{T}$ are given by the formula

$$\beta = \frac{d}{d\tau}\gamma(\tau)\Big|_{\tau=0} = \sum_{n\geq 1} \frac{(-1)^{n+1}}{n}(\alpha-\mathbb{1})^{*n}, \tag{3.11}$$

or equivalently,

$$\beta_u = \frac{d}{d\tau}\Big|_{\tau=0}\gamma(\tau)_u = \sum_{m=1}^{|u|} \frac{(-1)^{m+1}}{m} \sum_{(V_m\succ\cdots\succ V_1)\subset U} \alpha_{V_1}\cdots\alpha_{V_m}, \tag{3.12}$$

where in the inner summation only non-empty posets $V_1,\ldots,V_m$ are considered. For instance,

$$\begin{aligned} \beta_{\bullet} &= \alpha_{\bullet}, \\ \beta_{[\bullet]} &= \alpha_{[\bullet]} - \frac{1}{2}\alpha_{\bullet}^2, \\ \beta_{[[\bullet]]} &= \alpha_{[[\bullet]]} - \alpha_{\bullet}\,\alpha_{[\bullet]} + \frac{1}{3}\alpha_{\bullet}^3, \\ \beta_{[\bullet\bullet]} &= \alpha_{[\bullet\bullet]} - \alpha_{\bullet}\,\alpha_{[\bullet]} + \frac{1}{6}\alpha_{\bullet}^3. \end{aligned}$$

We thus have that $\beta = \log(\alpha)$, or equivalently, $\alpha = \exp(\beta)$, where the mutually reciprocal operators log and exp are defined as follows:

Definition 3.3 *Given* $\alpha \in \mathbb{R}^{\mathcal{F}}$ *with* $\alpha_e = 1$, $\log(\alpha) \in \mathbb{R}^{\mathcal{F}}$ *is defined as*

$$\log(\alpha) = \sum_{n\geq 1} \frac{(-1)^{n+1}}{n} (\alpha - \mathbb{1})^{*n}.$$

Given $\beta \in \mathbb{R}^{\mathcal{F}}$ *with* $\beta_e = 0$, $\exp(\beta) \in \mathbb{R}^{\mathcal{F}}$ *is defined as*

$$\exp(\beta) = \mathbb{1} + \sum_{n\geq 1} \frac{1}{n!} \beta^{*n}.$$

Though defined as infinite series, $\log(\alpha)$ and $\exp(\beta)$ are well defined as functions in $\mathbb{R}^{\mathcal{F}}$, because $\alpha_e = \mathbb{1}$ (resp. $\beta_e = 0$) implies that $((\alpha_e - \mathbb{1})^{*n})_u = 0$ (resp. $(\beta^{*n})_u = 0$) whenever $|u| < n$. (Property P2 in Subsection 2.5.)

Summing up, we have seen that if the formal solution $\tilde{y}(t)$ of a modified equation (3.5) satisfies (3.6), then the right-hand side of (3.5) is given as the B-series (3.10) where $\beta = \log(\alpha)$. We will next show that the solution $\tilde{y}(t)$ of such a modified equation actually satisfies (3.6): Substitution of $\tilde{y}(t) = B_{hf}(\gamma(t/h), y_0)$ in (3.5) and application of Theorem 3.2 gives

$$\frac{d}{dt} B_{hf}(\gamma(t/h), y_0) = \frac{1}{h} B_{hf}(\gamma(t/h) * \beta, y_0), \quad B_{hf}(\gamma(0), y_0) = y_0,$$

or equivalently,

$$B_{hf}(\frac{d}{d\tau}\gamma(\tau), y_0) = B_{hf}(\gamma(\tau) * \beta, y_0), \quad B_{hf}(\gamma(0), y_0) = y_0,$$

which will hold true if $\gamma(\tau)$ satisfies the ODE

$$\frac{d}{d\tau}\gamma(\tau) = \gamma(\tau) * \beta, \quad \gamma(0) = \mathbb{1}, \tag{3.13}$$

whose solution is $\gamma(\tau) = \exp(\tau\beta)$. Hence, $\tilde{y}(t) = B_{hf}(\gamma(t/h), y_0)$ with $\gamma(\tau) = \exp(\tau \log(\alpha))$ (which can be expanded as (3.9)) is the solution of the modified equation (3.5) with (3.10) and $\beta = \log(\alpha)$. We finally have that (3.6) holds true as $\exp(n \log(\alpha)) = \alpha^{*n}$ for all $n \in \mathbb{Z}$.

Obviously, the right-hand side of (1.1) can be written as $h^{-1}B_{hf}(\delta^\bullet, y)$, where $\delta^\bullet \in \mathfrak{g}$ is defined as $\delta^\bullet_u = 1$ if $u = \bullet$ and $\delta^\bullet_u = 0$ otherwise. Thus the exact solutions $y(t)$ of (1.1) admits the expansion $y(t) = B_{hf}(\gamma(t/h), y(0))$, where $\gamma(\tau) = \exp(\tau\,\delta^\bullet)$ is the solution of the ODE (3.13) with $\beta = \delta^\bullet$, that is, for each $u \in \mathcal{T}$ with $u = [v]$, $v \in \mathcal{F}$,

$$\frac{d}{d\tau}\gamma(\tau)_u = \gamma(\tau)_v, \quad \gamma(0)_u = 0,$$

which leads to $\gamma(\tau)_u = \exp(\tau\delta^\bullet)_u = \tau^{|u|}/u!$, where the factorial $u!$ of the rooted tree u is defined as in Theorem 2.3.

The fact that $\beta_\bullet = \alpha_\bullet$ implies that (3.5) will be a perturbation of size $\mathcal{O}(h)$ of the original ODE (1.1) provided that $\alpha_\bullet = 1$. More generally, (3.5) will be a perturbation of size $\mathcal{O}(h^r)$ of the original ODE (1.1) if and only if $\alpha_u = \exp(\delta^\bullet)_u = 1/u!$ for all $u \in \mathcal{T}$ with $|u| \leq r$, that is, if $\alpha \in \mathcal{G}$ corresponds to a method of order r for the original ODE (1.1).

Unfortunately, the series in powers of h defining the right-hand side (3.10) of the modified ODE (3.5) is in general divergent, even when the B-series expansion (3.4) of the integration method is convergent (for real analytic f, and sufficiently small h). For rigorous results based on modified equations, one has to consider a truncated version, and estimate the differences between the numerical solution and the solution of the truncated modified ODE.

3.3 Series of linear differential operators

Let $\psi_{hf}(y)$ be one step of an integration method applied to the ODE (1.1) that admits a B-series expansion $\psi_{hf}(y) = B_{hf}(\alpha, y)$ for some $\alpha \in \mathcal{G}$. Assume that we want to expand the composition of ψ_{hf} with a smooth function $g \in C^\infty(\mathbb{R}^d)$, that is, we want the expansion in powers of h of $g(B_{hf}(\alpha, y))$. Proceeding as in the proof of Proposition 2.10 when showing that (2.6) implies that $f(\psi_{hf}(y))$ admits the expansion (2.7), one readily gets that $g(B_{hf}(\alpha, y))$ can be expanded as

$$g(y) + \sum_{v=u_1\cdots u_m \in \mathcal{F}\backslash\{e\}} \frac{h^{|v|}}{\sigma(v)} \alpha_v \, g^{(m)}(y)(F_{u_1}(y), \ldots, F_{u_m}(y)), \tag{3.14}$$

where the summation goes over all non-empty forests $v \in \mathcal{F}\backslash\{e\}$, and for each such forest, $v = u_1 \cdots u_m$ is its unique decomposition in rooted trees $u_1, \ldots, u_m \in \mathcal{T}$.

This motivates us to make the following definitions.

Definition 3.4 *Given a smooth vector field f, the* elementary differential operator *X_u associated to a forest $u \in \mathcal{F}$ is an operator acting on smooth functions $g \in C^\infty(\mathbb{R}^d)$ defined as follows: If $u = e$, then $X_e\, g = g$, and if $u = u_1 \cdots u_m$, where $u_1, \ldots, u_m \in \mathcal{T}$, then the function $X_u\, g$ is defined for each $y \in \mathbb{R}^d$ as*

$$X_u\, g(y) = g^{(m)}(y)(F_{u_1}(y), \ldots, F_{u_m}(y)),$$

where the elementary differentials F_v associated to rooted trees v are given in Definition 2.7.

In particular, $X_\bullet g(y) = g'(y)f(y)$, and thus $X_\bullet$ is the Lie derivative along the smooth vector field f.

Definition 3.5 *Given a smooth vector field f, to each $\alpha \in \mathbb{R}^{\mathcal{F}}$ we associate the formal series $S_{hf}(\alpha)$ of differential operators acting on smooth functions $g \in C^\infty(\mathbb{R}^d)$ given as*

$$S_{hf}(\alpha) = \sum_{u \in \mathcal{F}} \frac{h^{|u|}}{\sigma(u)} \alpha_u\, X_u.$$

With Definition 3.4, we clearly have that, if $u \in \mathcal{F}$, then the ith component of the elementary differential $F_{[u]}$ associated to the rooted tree $[u]$ can be obtained as the action of X_u on the ith component of f, that is, $F^i_{[u]} = X_u f^i$. (Alternatively, $F^i_{[u]} = X_{[u]} e_i$, where e_i denotes the ith coordinate function, that is, $e_i(y) = y^i$.) We thus have that, for each $\alpha \in \mathcal{G}$,

$$B_{hf}(\alpha, y) = y + h\, S_{hf}([\alpha]^*) f(y), \tag{3.15}$$

where $[\alpha]^* \in \mathbb{R}^{\mathcal{F}}$ is determined from the values α_u for $u \in \mathcal{T}$ as

$$([\alpha]^*)_v = \alpha_{[v]} \quad \text{for each} \quad v \in \mathcal{F}. \tag{3.16}$$

Observe that, if $\alpha \in \mathcal{G}_{RK}$, that is, if $\exists \mu \in \mathcal{RK}$ such that $\forall u \in \mathcal{F}$, $\alpha_u = u(\mu)$, then $([\alpha]^*)_u = [u](\mu)$.

In what follows, we will assume that the vector field f is fixed, and will often drop the subindex hf from the notation. From the discussion at the beginning of the subsection, we have the following basic result:

Proposition 3.6 *For each $\alpha \in \mathcal{G}$ and each $g \in C^\infty(\mathbb{R}^d)$,*

$$g(B(\alpha, y)) = S(\alpha) g(y).$$

We thus have that a series of differential operators $S(\alpha)$ for $\alpha \in \mathcal{G}$ can be identified with a B-series $B(\alpha, y)$ that represents a near-to-identity map (recall that $\alpha_e = 1$ if $\alpha \in \mathcal{G}$). From Proposition 3.6, we have that, for arbitrary $g_1, g_2 \in C^\infty(\mathbb{R}^d)$,

$$S(\alpha) g_1 g_2 = (S(\alpha) g_1)\,(S(\alpha) g_2) \quad \text{if} \quad \alpha \in \mathcal{G}. \tag{3.17}$$

We will next see that, while in the B-series context, $\alpha * \beta$ does not make sense for arbitrary $\alpha, \beta \in \mathbb{R}^{\mathcal{F}}$, it does in terms of series of differential operators. We will first state the following auxiliary result concerning the product $*$ in Definition 2.15.

Proposition 3.7 *Given $\alpha, \beta \in \mathbb{R}^{\mathcal{F}}$,*

$$[\alpha * \beta]^* = \beta_e\, [\alpha]^* + \alpha * [\beta]^*. \tag{3.18}$$

We omit the proof of Proposition 3.7 because of its similarity to the proof of Theorem 2.14. Observe that, in the case where $\alpha, \beta \in \mathcal{G}_{RK}$, (that is, $\exists \mu, \mu' \in \mathcal{RK}$ such that $\forall u \in \mathcal{F}$, $\alpha_u = u(\mu)$ and $\beta_u = u(\mu)$) (3.18) reduces to (2.11) (with $(\alpha * \beta)_u = u(\mu\mu')$ for all $u \in \mathcal{F}$).

Lemma 3.8 *For each $\alpha \in \mathbb{R}^{\mathcal{F}}$ and each $N \geq 1$, there exist $\alpha^1, \ldots, \alpha^{m_N} \in \mathcal{G}$ and $\lambda_1, \ldots, \lambda_{m_N} \in \mathbb{R}$ such that,*

$$\alpha_u = \sum_{j=1}^{m_N} \lambda_j\, \alpha^j_u, \quad \textit{for all} \;\; u \in \mathcal{F} \;\textit{with}\; |u| \leq N.$$

Theorem 3.9 *For each $\alpha, \beta \in \mathbb{R}^{\mathcal{F}}$,*

$$S(\alpha)S(\beta) = S(\alpha * \beta). \tag{3.19}$$

Proof We will prove by induction on $N \geq 1$ that, for all $(\alpha, \beta) \in \mathbb{R}^{\mathcal{F}} \times \mathbb{R}^{\mathcal{F}}$,

$$S(\alpha)S(\beta)g - S(\alpha * \beta)g = \mathcal{O}(h^{N+1}). \tag{3.20}$$

By virtue of Lemma 3.8, and by bilinearity of both $*$ and the composition of linear differential operators, it is then enough to prove (3.20) for $\alpha, \beta \in \mathcal{G}$. According to Proposition 3.6, if $\alpha, \beta \in \mathcal{G}$, then

$$S(\alpha)S(\beta)g(y) - S(\alpha * \beta)g(y) = g(B(\beta, B(\alpha, y))) - g(B(\alpha * \beta, y)), \tag{3.21}$$

and thus we have to prove that, for all $(\alpha, \beta) \in \mathcal{G} \times \mathcal{G}$,

$$B(\beta, B(\alpha, y)) - B(\alpha * \beta, y) = \mathcal{O}(h^{N+1}),$$

but according to (3.15) and Proposition 3.6,

$$\begin{aligned} B(\beta, B(\alpha, y)) - B(\alpha * \beta, y) &= S(\alpha)B(\beta, y) - y - h\,S([\alpha * \beta]^*)f(y) \\ &= h\,S([\alpha]^*)f(y) + h\,S(\alpha)S([\beta]^*)f(y) - h\,S([\alpha * \beta]^*)f(y). \end{aligned}$$

By induction hypothesis, $S(\alpha)S([\beta]^*) = S(\alpha * [\beta]^*) + \mathcal{O}(h^N)$, and thus,

$$B(\beta, B(\alpha, y)) - B(\alpha * \beta, y) = h\,S([\alpha]^* + \alpha * [\beta]^* - [\alpha * \beta]^*)f(y) + \mathcal{O}(h^{N+1})$$

and the required result is obtained by virtue of Proposition 3.7. ■

We thus have that (3.5) defines an algebra homomorphism from the algebra $\mathbb{R}^{\mathcal{F}}$ to the algebra $\mathcal{E}[[h]]$, where $\mathcal{E}$ denotes the algebra of endomorphisms of $C^{\infty}(\mathbb{R}^d)$. In particular, $S_{hf}(e)$ is the identity operator. We also have that $S_{hf}(\exp(\beta)) = \exp(S_{hf}(\beta))$ (resp. $S_{hf}(\log(\alpha)) = \log(S_{hf}(\alpha))$) if $\beta_e = 0$ (resp. if $\alpha_e = 1$).

The following is a consequence of (3.17) and Lemma 3.8.

Proposition 3.10 *For each $\alpha \in \mathbb{R}^{\mathcal{F}}$,*

$$S(\alpha)g_1 g_2 = \sum_{(u,v) \in \mathcal{F} \times \mathcal{F}} \frac{h^{|u|+|v|}}{\sigma(u)\sigma(v)} \alpha(uv)\,(X_u g_1)\,(X_v g_2).$$

3.4 The Lie algebra of the Butcher group

We next introduce the Lie algebra $\mathfrak{g} = \log(\mathcal{G})$ of the Butcher group $\mathcal{G}$.

Proposition 3.11 *Given $\beta \in \mathbb{R}^{\mathcal{F}}$, the following statements are equivalent:*

- $\exp(\beta) \in \mathcal{G}$.
- *For all $u, v \in \mathcal{F}$,*

$$\beta_{uv} = \mathbb{1}_u \beta_v + \beta_u \mathbb{1}_v. \tag{3.22}$$

- *there exists a curve $\gamma : \mathbb{R} \to \mathcal{G}$ such that*

$$\gamma(0) = \mathbb{1}, \quad \frac{d}{dt}\gamma(t)\bigg|_{t=0} = \beta. \tag{3.23}$$

Clearly, (3.22) is equivalent to

$$\beta_e = 0, \quad \beta_{u_1 \cdots u_m} = 0, \quad \text{if } u_1, \ldots, u_m \in \mathcal{T}. \tag{3.24}$$

Hence, as in the case of $\mathcal{G}$, each $\beta \in \mathfrak{g}$ is uniquely determined by its values β_u for $u \in \mathcal{T}$.

Given a vector field f, for each $\beta \in \mathfrak{g}$, $S_{hf}(\beta)$ is a derivation of the algebra $C^\infty(\mathbb{R}^d)$, i.e.,

$$S_{hf}(\beta) g_1 g_2 = (S_{hf}(\beta) g_1)\, g_2 + g_1\, (S_{hf}(\beta) g_2) \quad \text{if} \quad \beta \in \mathfrak{g}. \tag{3.25}$$

That is, $S_{hf}(\beta)$ is (the Lie operator of) a vector field. The equality (3.25) follows from (3.17) by considering a curve $\gamma(\tau)$ in $\mathcal{G}$ satisfying (3.23). This can also be shown directly from the definition of $S_{hf}(\beta)$ by taking into account that, if $\beta \in \mathfrak{g}$, then $\beta_u = 0$ for all $u \in \mathcal{F}\backslash\mathcal{T}$, so that

$$S_{hf}(\beta) g(y) = \frac{\partial}{\partial y} g(y) B_{hf}(\beta, y).$$

Proposition 3.11 implies that $\mathfrak{g}$ is a Lie subalgebra of the Lie algebra $\mathbb{R}^{\mathcal{F}}$ with bracket $[\alpha, \beta] = \alpha * \beta - \beta * \alpha$. Thus, the restriction of S_{hf} to $\mathfrak{g}$ defines a homomorphism from the Lie algebra $\mathfrak{g}$ to the Lie algebra of (formal series of) vector fields.

3.5 The pre-Lie algebra structure on $\mathfrak{g}$

We next show that the Lie algebra $\mathfrak{g}$ has a pre-Lie algebra structure. A nice exposition on pre-Lie algebras can be found in [?].

It will be useful to consider the projection $\pi : \mathbb{R}^{\mathcal{F}} \to \mathfrak{g}$ given by $\pi(\beta)_u = 0$ for each $u \in \mathcal{F}\backslash\mathcal{T}$, and $\pi(\beta)_u = \beta_u$ for each $u \in \mathcal{T}$. Observe that the restriction of π to $\mathcal{G}$ is (as the restriction to $\mathcal{G}$ of the logarithm) a bijection from $\mathcal{G}$ to $\mathfrak{g}$.

Property P1 in Subsection 2.5 implies that $\mathfrak{g}$ has a left $\mathcal{H}^*$-module structure, with $\mathcal{H}^* \otimes \mathfrak{g} \to \mathfrak{g}$ defined as follows.

Definition 3.12 *Given $\alpha \in \mathcal{H}^*$ and $\beta \in \mathfrak{g}$,*

$$\alpha \cdot \beta := \pi(\alpha * \beta).$$

We use the symbol $\rhd$ to denote the restriction to $\mathfrak{g} \otimes \mathfrak{g}$ of that $\mathcal{H}^*$-module map. Thus, given $\beta, \beta', \beta'' \in \mathfrak{g}$,

$$\beta \rhd (\beta' \rhd \beta'') = \beta \cdot (\beta' \cdot \beta'') = (\beta * \beta') \cdot \beta''. \tag{3.26}$$

Since $\mathfrak{g}$ is a Lie algebra under the bracket $[\beta, \beta'] = \beta * \beta' - \beta' * \beta$, defined for $\beta, \beta' \in \mathfrak{g}$, we have that $\pi([\beta, \beta']) = [\beta, \beta']$, and thus

$$\forall \beta, \beta' \in \mathfrak{g}, \quad \beta \rhd \beta' - \beta' \rhd \beta = [\beta, \beta']. \tag{3.27}$$

Proposition 3.13 *The vector space $\mathfrak{g}$ with the restriction to $\mathfrak{g}\otimes\mathfrak{g}$ of $\rhd$ has a left pre-Lie algebra structure.*

Proof *According to the definition of left pre-Lie algebra, we have to prove that, given $\beta,\beta',\beta''\in\mathfrak{g}$,*

$$(\beta\rhd\beta'-\beta'\rhd\beta)\rhd\beta'' \;=\; \beta'\rhd(\beta\rhd\beta'')-(\beta'\rhd\beta)\rhd\beta''.$$

Indeed, given $\beta,\beta',\beta''\in\mathfrak{g}$, taking into acount (3.27) and (3.26),

$$\begin{aligned}(\beta\rhd\beta'-\beta'\rhd\beta)\rhd\beta'' &= [\beta,\beta']\rhd\beta''\\ &= (\beta'*\beta-\beta*\beta')\cdot\beta''\\ &= (\beta'*\beta)\cdot\beta''-(\beta*\beta')\cdot\beta''\\ &= \beta'\rhd(\beta\rhd\beta'')-\beta\rhd(\beta'\rhd\beta'').\end{aligned}$$

■

Proposition 3.14 *If $\alpha\in\mathcal{H}^*$ and $\beta\in\mathfrak{g}$, then*

$$B(\alpha\cdot\beta,y)=S(\alpha)B(\beta,y).$$

In particular, if $\alpha,\beta\in\mathfrak{g}$, then $B(\alpha\rhd\beta,y)=\frac{\partial}{\partial y}B(\beta,y)\cdot B(\alpha,y)$.

Proof It follows from Theorem 3.9, since for each $\gamma\in\mathcal{H}^*$, the ith compoment of $B(\gamma)=B(\pi(\gamma))$ coincides with the action of $S(\gamma)$ on the ith coordinate function $e_i\in C^\infty(\mathbb{R}^d)$, the required result is equivalent to

$$S(\alpha*\beta)e_i=S(\alpha)S(\beta)e_i,\quad i=1,\ldots,d.$$

■

This shows that, given a smooth vector field f, the definition of the B-series $B_{hf}(\beta,\cdot)$ for each $\beta\in\mathfrak{g}$ determines a homomorphism from the left pre-Lie algebra $\mathfrak{g}$ to the left pre-Lie algebra of (formal series of) smooth vector fields in $\mathbb{R}^d$: Recall that the binary operation $\rhd$ endowing the vector space of smooth vector fields in $\mathbb{R}^d$ (given as maps $f:\mathbb{R}^d\to\mathbb{R}^d$) with a left pre-Lie algebra structure is defined as $(f\rhd g)(y)=\frac{\partial}{\partial y}g(y)\cdot f(y)$, that is, $f\rhd g$ is the action on g of the Lie operator associated to f.

The next result gives a characterization of Lie algebras that are obtained in this way from a left pre-Lie algebra structure.

Proposition 3.15 *Given a Lie algebra $\mathcal{L}$, consider its universal enveloping algebra $\mathcal{U}(\mathcal{L})$. The vector space $\mathcal{L}$ can be endowed with a left pre-Lie algebra operation $\rhd:\mathcal{L}\otimes\mathcal{L}\to\mathcal{L}$ such that for all $\ell,\ell'\in\mathcal{L}$,*

$$[\ell,\ell']=\ell\rhd\ell'-\ell'\rhd\ell$$

if and only if $\mathcal{L}$ has a left $\mathcal{U}(\mathcal{L})$-module structure. In that case, $\rhd$ is the restriction to $\mathcal{L}\otimes\mathcal{L}$ of the left-module map $\mathcal{U}(\mathcal{L})\otimes\mathcal{L}\to\mathcal{L}$.

Proof Clearly, the existence of a left $\mathcal{U}(\mathcal{L})$-module structure of $\mathcal{L}$ can be replaced in the statement of Proposition 3.15 by the existence of a triplet $(\mathcal{A}, \nu, \cdot)$, where $\mathcal{A}$ is an associative algebra, $\nu : \mathcal{L} \to \mathcal{A}$ is a monomorphism of Lie algebras, and $\cdot : \mathcal{A} \otimes \mathcal{L} \to \mathcal{L}$ is a left $\mathcal{A}$-module structure map. The 'if' part of the statement can be proven as Proposition 3.13. The 'only if' part is obtained by considering $\mathcal{A}$ as the algebra of endomorphisms of the left pre-Lie algebra $\mathcal{L}$, the canonical left $\mathcal{A}$-module structure of $\mathcal{L}$, and the injection $\nu : \mathcal{L} \to \mathcal{A}$ defined as $\nu(\ell) \cdot \ell' = \ell \rhd \ell'$, and proving that ν is a Lie algebra morphism. Indeed, by definition of ν and the left pre-Lie algebra condition, we have that

$$\begin{aligned} \nu([\ell, \ell']) \cdot \ell'' &= (\ell \rhd \ell' - \ell' \rhd \ell) \rhd \ell'' \\ &= \ell \rhd (\ell' \rhd \ell'') - \ell' \rhd (\ell \rhd \ell'') \\ &= \nu(\ell) \cdot (\nu(\ell') \cdot \ell'') - \nu(\ell') \cdot (\nu(\ell) \cdot \ell'') \\ &= [\nu(\ell), \nu(\ell')] \cdot \ell''. \end{aligned}$$

■

Let us consider the vector space $\mathfrak{g}_0$ of elements in $\mathfrak{g}$ with finite support, or equivalently,

$$\mathfrak{g}_0 = \{\beta \in \mathfrak{g} \ : \ \exists N \geq 1 \text{ such that } \beta_u = 0 \text{ if } |u| \geq N\}. \tag{3.28}$$

Clearly, $\mathfrak{g}_0$ is a left pre-Lie subalgebra of $\mathfrak{g}$. The set $\mathcal{T}$ of rooted trees can be identified with the basis $\{\delta^u \ : \ u \in \mathcal{T}\}$ of $\mathfrak{g}_0$, where for each $u \in \mathcal{T}$,

$$(\delta^u)_v = \begin{cases} \sigma(u) & \text{if } v = u \\ 0 & \text{otherwise,} \end{cases} \tag{3.29}$$

where the normalization factor $\sigma(u)$ is determined for each $u \in \mathcal{T}$ in Definition 2.9. We thus have that, for each $\beta \in \mathfrak{g}_0$,

$$\beta = \sum_{u \in \mathcal{T}} \frac{\beta_u}{\sigma(u)} \, \delta^u.$$

A left pre-Lie algebra structure is determined for the set $\mathbb{R}\mathcal{T}$ of linear combinations or rooted trees by identifying each δ^u with the rooted tree u. It can be shown that the pre-Lie algebra $\mathfrak{g}_0$ is generated by $\delta^\bullet$ and also, that

$$\delta^u \rhd \delta^v = \sum_{i=1}^{|v|} \delta^{w_i}$$

where each w_i is the rooted tree obtained by grafting the rooted tree u to the ith vertex of v. This is precisely the left pre-Lie algebra of rooted trees as introduced by Chapoton [8], who proves the following fundamental result:

Theorem 3.16 *The left pre-Lie algebra $\mathfrak{g}_0$ of rooted trees is the free left pre-Lie algebra in one generator.*

In Subsection 4.3 below, an alternative to the original proof in [8] is given in Theorem 4.3.

3.6 Bibliographical comments

A direct proof of Theorem 3.2 based on the Fàa di Bruno formula was given in the original paper [19]. The construction of modified equation for general B-series methods is originally due to Hairer [20]. The formula (3.12) for the coefficients of the modified equations is closely related to the one originally introduced in [20]. An alternative approach to recursively compute the coefficients $\beta(u)$ and $\gamma(u)$ from the ODE (3.13) supplemented by the condition $\gamma(1) = \alpha$ was considered in [5], which only makes use of the formula (2.18) of composition of B-series. The interpolatory argument adopted in Subsection 3.2 was introduced in [26], and explicit expressions for $\log(\alpha)_u$ in terms of rooted trees are given in [27] and [9]. The linear differential operators X_u were first considered in the context of Runge-Kutta methods by Merson [25]. They can also be found later on in [17], [26, 27]. In [12] and [11] series of such differential operators are successfully applied to study preservation properties of integration methods that can be expanded as B-series. In [8], the left pre-Lie algebra structure on the linear span of rooted trees is studied, and it is shown that it is the free left pre-Lie algebra in one generator (see Theorem 3.16 below).

4 Hopf algebras of rooted trees

Clearly, all the material in Subsection 2.5 would be valid if the set $\mathbb{R}^{\mathcal{F}}$ of functions had been replaced by the set $\mathcal{R}^{\mathcal{F}}$ of maps from $\mathcal{F}$ to an arbitrary commutative algebra $\mathcal{R}$. In particular, this gives a group structure $\mathbb{G}(\mathcal{R})$ on the set of maps $\alpha : \mathcal{F} \to \mathcal{R}$ satisfying that $\alpha_{uv} = \alpha_u \alpha_v$. In this sense, the Butcher group is actually an *affine group scheme* $\mathbb{G}(\cdot)$, a functor from commutative algebras to groups. The category of group schemes is equivalent to the category of commutative Hopf algebras [**?**]: Given a group scheme $\mathbb{G}(\cdot)$, there exists a commutative Hopf algebra $\mathcal{H}$ such that $\mathbb{G}(\mathcal{R})$ is the group of algebra maps from $\mathcal{H}$ to $\mathcal{R}$ with the convolution product as group law. The coproduct in $\mathcal{H}$ can be obtained by dualizing the group law of the affine group scheme $\mathbb{G}(\cdot)$. Applying this to the particular case of Butcher's affine group scheme gives rise to a commutative Hopf algebra $\mathcal{H}$ on the vector space of linear combinations of forests of rooted trees.

As an alternative construction of $\mathcal{H}$, the algebra $\mathbb{R}[\mathcal{T}]$ of functions on $\mathcal{RK}$ generated by $\mathcal{T}$ (each $u \in \mathcal{T}$ viewed as a function in $\mathbb{R}^{\mathcal{RK}}$ as defined in Section 2) can be endowed with a commutative Hopf algebra structure derived from the semigroup structure of the set of RK methods. This can be done by observing that the results in Subsection 2.4 imply that the algebra $\mathbb{R}[\mathcal{T}]$ is a subbialgebra of the bialgebra of representative functions [21] of the semigroup $\mathcal{RK}$. (In particular, the coassociativity of the coproduct in $\mathcal{H}$ comes as a direct consequence of the associativity of the composition of RK methods.) The existence of the antipode is guaranteed, in this context, by the conilpotency of the coproduct.

There is another important Hopf algebra associated to the Butcher group

$\mathcal{G}$, the dual Hopf algebra [28] $\mathcal{H}^\circ$ of $\mathcal{H}$. The Butcher group $\mathcal{G}$ coincides with the group $G(\mathcal{H}^\circ)$ of *group-like elements* of $\mathcal{H}^\circ$. The Lie algebra $\mathfrak{g}$ of the group $\mathcal{G}$ is the Lie algebra $P(\mathcal{H}^\circ)$ of *primitive elements* of $\mathcal{H}^\circ$.

4.1 The commutative Hopf algebra of rooted trees

The preceeding discussion leads to a commutative Hopf algebra structure on the vector space spanned by the set $\mathcal{F}$ of forests. As an algebra, it is just the algebra $\mathbb{R}[\mathcal{T}]$ of polynomials in the commuting indeterminates $u \in \mathcal{T}$ (with the empty forest e as the unit), or equivalently, the symmetric algebra $S(\mathcal{V})$ over the vector space $\mathcal{V}$ spanned by the set of rooted trees $\mathcal{T}$. The bijection $\mathcal{F} \to \mathcal{T}$ that gives the rooted tree $[u]$ for each forest u induces an isomorphism of vector spaces $B : S(\mathcal{V}) \to \mathcal{V}$, where $B(u) = [u]$ for each $u \in \mathcal{F}$.

The coproduct

$$\begin{array}{rcl} \Delta : \mathcal{H} & \to & \mathcal{H} \otimes \mathcal{H} \\ u & \mapsto & \Delta(u) \end{array}$$

is defined by determining $\Delta(u)$ for each $u \in \mathcal{F}$ as follows: If U is a partially ordered set representing a forest $u \in \mathcal{F}$,

$$\begin{aligned} \Delta(u) &= \sum_{(V \succ W) \subset U} v \otimes w \\ &= u \otimes e + e \otimes u + \sum_{\substack{(V \succ W) \subset U \\ V, W \neq \emptyset}} v \otimes w, \end{aligned} \tag{4.1}$$

where V and W represent the forests v and w respectively. The coassociativity (which can be seen as a consequence of the associativity of the composition of RK methods), can be directly checked from the definition of $\Delta(u)$ for $u \in \mathcal{F}$ (exactly like the associativity of the group law of $\mathcal{G}$): For each $u \in \mathcal{F}$,

$$(\Delta \otimes \mathrm{id}_{\mathcal{H}}) \circ \Delta(u) = (\mathrm{id}_{\mathcal{H}} \otimes \Delta) \circ \Delta(u) = \sum_{(V_1 \succ V_2 \succ V_3) \subset U} v_1 \otimes v_2 \otimes v_3.$$

The counit $\mathbb{1} : \mathcal{H} \to \mathbb{R}$ is defined by setting $\mathbb{1}(e) = 1$ and $\mathbb{1}(u) = 0$ for all $u \in \mathcal{F} \backslash \{e\}$. The antipode $s(u)$ can be defined for each $u \in \mathcal{H}$ by

$$\begin{aligned} s(u) &= -u - \sum_{\substack{(V \succ W) \subset U \\ V, W \neq \emptyset}} v \otimes s(w) \\ &= -u + \sum_{m \geq 2} (-1)^m \sum_{\substack{(V1 \succ V_2 \succ V_m) \subset U \\ V_1, \dots, V_m \neq \emptyset}} v_1 \otimes \cdots \otimes v_m. \end{aligned}$$

It is obvious from the definition of the coproduct of Δ that, as a coalgebra, $\mathcal{H}$ is conilpotent (i.e., connected), that is to say, for each $u \in \mathcal{H}$ such that $\mathbb{1}(u) = 0$, there exists an $n \geq 0$ such that $\overline{\Delta}^n(u) = 0$, where $\overline{\Delta}(u) = \Delta(u) - u \otimes e - e \otimes u$, and $\overline{\Delta}^{n+1} = (\overline{\Delta}^n \otimes \mathrm{id}_{\mathcal{H}}) \circ \overline{\Delta}$.

Clearly, the Hopf algebra $\mathcal{H}$ is compatible with the grading $\mathcal{H} = \bigoplus_{n\geq 0} \mathcal{H}_n$, where each homogeneous component $\mathcal{H}_n$ is spanned by the set $\mathcal{F}_n$ of forests with n vertices. Thus, $\mathcal{H}$ has a graded connected Hopf algebra structure.

Next result is a dual version of (2.12) (and also of (3.18)).

Proposition 4.1 (Kreimer) *Consider the linear map $L : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H} \otimes \mathcal{V}$ defined as $L(u \otimes v) = u \otimes [v]$ for each $u, v \in \mathcal{F}$. Then, for each $u \in \mathcal{H}$,*

$$\Delta([u]) = [u] \otimes e + L(\Delta(u)) \tag{4.2}$$

Proof If U' is the partially ordered set representing $[u]$, and $x \in U'$ is the root of U', then the partially ordered set $U = U' \backslash \{x\}$ represents u, and the pairs of partially ordered sets $(V', W') \neq (U, \emptyset)$ such that $(V' \succ W') \subset U'$ are in one-to-one correspondence with the pairs (V, W) such that $(V \succ W) \subset U$, with $V = V'$ and $W = W' \backslash \{x\}$. This implies that

$$\Delta([u]) = \sum_{(V \succ W') \subset U'} v \otimes w' = [u] \otimes e + \sum_{(V \succ W) \subset U} v \otimes [w].$$

■

4.2 The dual algebra $\mathcal{H}^*$ and the dual Hopf algebra $\mathcal{H}^\circ$

Given a linear form $\alpha \in \mathcal{H}^*$, we denote as $\langle \alpha, u \rangle$ the value of α at $u \in \mathcal{H}$. When $u \in \mathcal{F}$, we will keep the notation $\langle \alpha, u \rangle = \alpha_u$. The coalgebra structure of $\mathcal{H}$ induces an associative unital algebra structure on the algebraic dual $\mathcal{H}^*$ of $\mathcal{H}$, with the multiplication $*$ given by

$$\langle \alpha * \beta, u \rangle = \langle \alpha \otimes \beta, \Delta(u) \rangle,$$

where, as usual, $\mathcal{H}^* \otimes \mathcal{H}^*$ is considered as a subspace of $(\mathcal{H} \otimes \mathcal{H})^*$ (that is, given $\alpha, \beta \in \mathcal{H}^*$, $\alpha \otimes \beta \in (\mathcal{H} \otimes \mathcal{H})^*$ is defined as $\langle \alpha \otimes \beta, u \otimes v \rangle = \langle \alpha, u \rangle \langle \beta, v \rangle$).

Clearly, the algebra $\mathcal{H}^*$ coincides with the algebra structure on $\mathbb{R}^{\mathcal{F}}$ defined in Subsection 2.5.

Given $\alpha \in \mathcal{H}^*$, consider $\Delta(\alpha) \in (\mathcal{H} \otimes \mathcal{H})^*$.

$$\langle \Delta(\alpha), u \otimes v \rangle = \langle \alpha, uv \rangle.$$

Let $\mathcal{H}^\circ$ be the subalgebra of $\mathcal{H}^*$ consisting of $\alpha \in \mathcal{H}^*$ such that $\Delta(\alpha) \in \mathcal{H}^* \otimes \mathcal{H}^*$. Following Sweedler [28], we have that $\mathcal{H}^\circ$ has a cocommutative Hopf algebra structure. The group of group-like elements

$$G(\mathcal{H}^\circ) = \{\alpha \in \mathcal{H}^\circ \ : \ \Delta(\alpha) = \alpha \otimes \alpha\}$$

of $\mathcal{H}^\circ$ (or equivalently, the group of *characters* of $\mathcal{H}$) coincides with the Butcher group $\mathcal{G}$. The Lie algebra of primitive elements

$$P(\mathcal{H}^\circ) = \{\beta \in \mathcal{H}^\circ \ : \ \Delta(\beta) = \beta \otimes \mathbb{1} + \mathbb{1} \otimes \beta\}$$

of $\mathcal{H}^\circ$ (or equivalently, the Lie algebra of *infinitesimal characters* of $\mathcal{H}$) coincides with the Lie algebra $\mathfrak{g}$ of the Butcher group, which as shown in Subsection 3.5, has actually a left pre-Lie algebra structure.

Observe that the notation introduced in (3.16) corresponds to the dualization $[\cdot]^* : \mathcal{H}^* \to \mathcal{H}^*$ of $[\cdot] : \mathcal{H} \to \mathcal{H}$ given by

$$\langle [\alpha]^*, u\rangle = \langle \alpha, [u]\rangle, \quad \text{for each} \quad u \in \mathcal{H}.$$

The Lie (and pre-Lie) subalgebra $\mathfrak{g}_0$ of $\mathfrak{g}$ considered in Subsection 3.5 is the Lie algebra of primitive elements of the graded dual Hopf algebra $\mathcal{H}^{\mathrm{gr}*} \subset \mathcal{H}^\circ$ of $\mathcal{H}$. The dual basis of the basis $\{\frac{1}{\sigma(u)} u \ : \ u \in \mathcal{F}\}$ of $\mathcal{H}$ is $\{\delta^u \in \mathcal{H}^* \ : \ u \in \mathcal{H}\}$, where each δ^u is defined by (3.29). Clearly, one has that,

$$\forall u \in \mathcal{F}, \quad [\delta^{[u]}]^* = \delta^u.$$

The following result can be proven from the definitions of δ^u, the product $*$, and the $\mathcal{H}^*$-module map $\cdot$.

Proposition 4.2 *For each $u \in \mathcal{F}$ and $v \in \mathcal{T}$,*

$$\begin{aligned} \delta^u \cdot \delta^\bullet &= \delta^{[u]}, \\ \delta^v * \delta^u &= \delta^{vu} + \sum_{i=1}^{|u|} \delta^{w_i}, \end{aligned}$$

where the forests w_i are obtained by grafting the rooted tree v to each of the vertices of the forest u.

By the Cartier-Milnor-Moore theorem, the Hopf algebra $\mathcal{H}^{\mathrm{gr}*}$ is isomorphic to the universal enveloping Hopf algebra $\mathcal{U}(\mathfrak{g}_0)$ of the Lie algebra $\mathfrak{g}_0$ (defined in (3.28)).

4.3 B-series and series of differential operators revisited

Recall that, given a left pre-Lie algebra $\mathcal{P}$ with pre-Lie operation $\rhd$, Proposition 3.15 guarantees the existence of a left $\mathcal{U}(\mathcal{P})$-module structure on $\mathcal{P}$ such that, for each $p, p' \in \mathcal{P}$, $p \rhd p' = \nu(p) \cdot p'$, where ν is the canonical monomorphism of Lie algebras from $\mathcal{P}$ to $\mathcal{U}(\mathcal{P})$.

Theorem 4.3 *Given a left pre-Lie algebra $\mathcal{P}$, for each element $p \in \mathcal{P}$, there exists a unique pre-Lie algebra morphism $B_p : \mathfrak{g}_0 \to \mathcal{P}$ and a unique Hopf algebra morphism $S_p : \mathcal{U}(\mathfrak{g}_0) \to \mathcal{U}(\mathcal{P})$ such that $B_p(\delta^\bullet) = p$ and for each $\alpha \in \mathcal{H}^{\mathrm{gr}*}$, $\beta \in \mathfrak{g}_0$,*

$$S_p(\alpha) \cdot B_p(\beta) = B_p(\alpha \cdot \beta), \quad \forall (\alpha, \beta) \in \mathcal{H}^{\mathrm{gr}*} \times \mathfrak{g}_0 \tag{4.3}$$

$$S_p(\beta) = \nu(B_p(\beta)), \quad \forall \beta \in \mathfrak{g}_0. \tag{4.4}$$

Furthermore, if $\mathcal{P}$ is a graded left pre-Lie algebra $\mathcal{P} = \bigoplus_{n\geq 1} \mathcal{P}_n$ (which induces a grading of the Hopf algebra $\mathcal{U}(\mathcal{P})$) and $p \in \mathcal{P}_1$, then S_p and B_p are compatible with the grading of $\mathcal{H}^{\mathrm{gr}*}$, $\mathcal{P}$, and $\mathcal{U}(\mathcal{P})$.

Proof If there exists a pre-Lie algebra morphism $B_p : \mathfrak{g}_0 \to \mathcal{P}$ such that $B_p(\delta^{\bullet}) = p$, then by the universal property of $\mathcal{H}^{\mathrm{gr}*} = \mathcal{U}(\mathfrak{g}_0)$, there exists a unique algebra morphism $S_p : \mathcal{U}(\mathfrak{g}_0) \to \mathcal{U}(\mathcal{P})$ such that (4.4). In order to prove (4.3), it is clearly enough to check it for $\alpha = \beta_1 * \cdots * \beta_r$, where $\beta_1, \ldots, \beta_r \in \mathfrak{g}_0$:

$$\begin{aligned} S_p(\alpha) \cdot B_p(\beta) &= S_p(\beta_1) \cdot (S_p(\beta_2) \cdots (S_p(\beta_r) \cdot B_p(\beta)) \cdots) \\ &= B_p(\beta_1) \rhd (B_p(\beta_2) \rhd \cdots (B_p(\beta_r) \rhd B_p(\beta)) \cdots) \\ &= B_p(\beta_1 \rhd (\beta_2 \rhd \cdots (\beta_r \rhd \beta) \cdots)) \\ &= B_p(\beta_1 \cdot (\beta_2 \cdot \cdots (\beta_r \cdot \beta) \cdots)) \\ &= B_p(\alpha \cdot \beta). \end{aligned}$$

If such B_p and S_p exist, then according to Proposition 4.2, for each $u \in \mathcal{F}$ and $v \in \mathcal{T}$,

$$B_p(\delta^{[u]}) = B_p(\delta^u \cdot \delta^{\bullet}) = S_p(\delta^u) \cdot p, \tag{4.5}$$

$$S_p(\delta^{vu}) = S_p(\delta^v) * S_p(\delta^u) - \sum_{i=1}^{|u|} S_p(\delta^{w_i}), \tag{4.6}$$

where the forests w_i (determined as in Proposition 4.2) have the same number m of connected components as the forest u. This allows to prove by induction on $m + |u|$ (where $u = u_1 \cdots u_m$, $u_1, \ldots, u_m \in \mathcal{T}$) that $S_p(\delta^u)$, $u \in \mathcal{F}$, are uniquely determined from (4.4)–(4.6). In turn, (4.5) determines uniquely $B_p(\delta^v)$, $v \in \mathcal{T}$. We thus have the uniqueness of S_p and B_p. We next show that S_p and B_p uniquely determined by (4.4)–(4.6) actually satisfy the required conditions.

Proposition 4.2 and (4.6) imply that S_p is an algebra morphism. This, together with (4.5) imply that (4.3) holds true. The map B_p being a pre-Lie algebra morphism is a consequence of (4.5), since for $\alpha, \beta \in \mathfrak{g}_0$,

$$\begin{aligned} B_p(\alpha \rhd \beta) &= B_p(\alpha \cdot \beta) \\ &= S_p(\alpha) \cdot B_p(\beta) \\ &= \nu(B_p(\alpha)) \cdot B_p(\beta) \\ &= B_p(\alpha) \rhd B_p(\beta). \end{aligned}$$

In order to prove that S_p is a Hopf algebra morphism, we have to check that $\mathcal{H}^{\mathrm{gr}*} \overset{S_p}{\to} \mathcal{U}(\mathcal{P}) \overset{\Delta}{\to} \mathcal{U}(\mathcal{P}) \otimes \mathcal{U}(\mathcal{P})$ and $\mathcal{H}^{\mathrm{gr}*} \overset{\Delta}{\to} \mathcal{H}^{\mathrm{gr}*} \otimes \mathcal{H}^{\mathrm{gr}*} \overset{S_p \otimes S_p}{\to} \mathcal{U}(\mathcal{P}) \otimes \mathcal{U}(\mathcal{P})$ coincide. But both are algebra morphisms from $\mathcal{H}^{\mathrm{gr}*}$ to $\mathcal{U}(\mathcal{P}) \otimes \mathcal{U}(\mathcal{P})$ whose restriction to $\mathfrak{g}_0$ coincide, and thus the universal property of $\mathcal{H}^{\mathrm{gr}*} = \mathcal{U}(\mathfrak{g}_0)$ gives the required result. It is straightforward to check that, in case $\mathcal{P}$ is a graded Lie algebra, B_p and S_p are compatible with the gradings of $\mathcal{H}^{\mathrm{gr}*}$, $\mathcal{P}$, and $\mathcal{U}(\mathcal{P})$. ∎

Theorem 4.4 *Given a left pre-Lie algebra $\mathcal{P}$ that is compatible with a decreasing filtration and complete with respect to it, let $\bar{\mathcal{U}}(\mathcal{P})$ be the complete universal enveloping algebra of the Lie algebra $\mathcal{P}$. Then for each element $p \in \mathcal{P}$, there* *exists a unique filtered pre-Lie algebra morphism $B_p : \mathfrak{g} \to \mathcal{P}$ and a unique filtered algebra morphism $S_p : \mathcal{H}^* \to \bar{\mathcal{U}}(\mathcal{P})$ such that $B_p(\delta^\bullet) = p$, and (4.3)–(4.4) holds for each $\alpha \in \mathcal{H}^*$, $\beta \in \mathfrak{g}$.*

As a consequence of the last statement in Theorem 4.4, we have that the restriction to $\mathcal{G} = G(\mathcal{H}^\circ)$ of S_p defines a morphism of topological groups from $\mathcal{G} = \exp(\mathfrak{g})$ to $\exp(\mathcal{P})$ (with the topology induced by their decreasing filtrations).

Consider the left pre-Lie algebra $\mathcal{P}$ of series of smooth vector fields in $\mathbb{R}^d$ of the form $hf_1 + h^2 f_2 + \cdots$, where each f_j is a smooth vector field in $\mathbb{R}^d$. The complete universal enveloping algebra $\bar{\mathcal{U}}(\mathcal{P})$ can be realized as the completion of the graded algebra of endomorphisms parametrized by h generated by derivations of the form $h^n\, g(y)\, \frac{\partial}{\partial y^j}$, where $n \geq 1$, $j \in \{1, \ldots, d\}$, and $g \in C^\infty(\mathbb{R}^d)$. We have seen in Subsection 3.5 that, given a smooth vector field f, the definition of B-series $B_{hf}(\beta) := B_{hf}(\beta, \cdot)$ for each $\beta \in \mathfrak{g}$ determines a morphism from the left pre-Lie algebra $\mathfrak{g}$ to the left pre-Lie algebra $\mathcal{P}$ satisfying that $B_{hf}(\delta^\bullet) = hf$. Theorem 4.4 implies that such a pre-Lie algebra morphism $B_{hf} : \mathfrak{g} \to \mathcal{P}$ satisfying $B_{hf}(\delta^\bullet) = hf$ is unique. We have also seen that the map $S_{hf} : \mathcal{H}^* \to \bar{U}(\mathcal{P})$ defined in Subsection 3.3 is an algebra morphism and satisfies (4.3)–(4.4), and thus coincides with the unique algebra map S_p, $p = hf$, given by Theorem 4.4.

Furthermore, if $\alpha \in \mathcal{H}^\circ$ with

$$\Delta(\alpha) = \sum_{j=1}^{m} \beta_j \otimes \gamma_j,$$

then, for each $(u, v) \in \mathcal{F} \times \mathcal{F}$,

$$\alpha_{uv} = \langle \alpha, uv \rangle = \langle \Delta\alpha, u \otimes v \rangle = \sum_{j=1}^{m} (\beta^{[j]})_u \otimes (\gamma^{[j]})_v,$$

and thus, by virtue of Proposition 3.10, we have that for all $g_1, g_2 \in C^\infty(\mathbb{R}^d)$,

$$S_{hf}(\alpha) g_1 g_2 = \sum_{j=1}^{m} \left(S_{hf}(\beta^{[j]}) g_1 \right) \left(S_{hf}(\gamma^{[j]}) g_2 \right).$$

4.4 A universal property of the commutative Hopf algebra of rooted trees

Consider the vector space $\mathcal{V} = \mathbb{R}\,\mathcal{T}$. The restriction to $\mathcal{V}$ of $\Delta - (\mathrm{id}_{\mathcal{H}} \otimes e)$ is a linear map $\rho : \mathcal{V} \to \mathcal{H} \otimes \mathcal{V}$. (The coassociativity of Δ implies that ρ is a left $\mathcal{H}$-comodule map.)

Definition 4.5 *Following [23], we say that a commutative Hopf algebra $\widetilde{\mathcal{H}}$ is a* combinatorial right-sided Hopf algebra *if the following conditions hold:*

- *As a coalgebra, $\widetilde{\mathcal{H}}$ is conilpotent.*
- *As an algebra, $\widetilde{\mathcal{H}}$ is freely generated by a subspace $\widetilde{\mathcal{V}}$ having a left $\widetilde{\mathcal{H}}$-comodule structure with comodule map $\tilde{\rho} : \widetilde{\mathcal{V}} \to \widetilde{\mathcal{H}} \otimes \widetilde{\mathcal{V}}$ satisfying that $\tilde{\Delta}(v) = v \otimes \tilde{e} + \tilde{\rho}(v)$.*

The following result is a dual version of Theorem 4.3

Proposition 4.6 *Let $\widetilde{\mathcal{H}}$ be a commutative combinatorial right-sided Hopf algebra and let $\widetilde{\mathcal{V}}$ be the corresponding vector space $\widetilde{\mathcal{V}}$. Given a linear form $\tilde{\eta} : \widetilde{\mathcal{V}} \to \mathbb{R}$, there exists a unique Hopf algebra morphism $\phi : \widetilde{\mathcal{H}} \to \mathcal{H}$ such that*

$$\phi(\widetilde{\mathcal{V}}) \subset \mathcal{V} \quad \text{and} \quad \tilde{\eta} = \eta \circ \phi,$$

where $\eta : \mathcal{V} \to \mathbb{R}$ is given by $\eta(\bullet) = 1$, $\eta(u) = 0$ if $|u| > 1$.

The following result allows constructing the Hopf algebra morphism ϕ in Proposition 4.6 in a recursive way.

Proposition 4.7 *Under the assumptions of Proposition 4.6, the following diagram commutes*

$$\begin{array}{ccc} \widetilde{\mathcal{V}} & \xrightarrow{\tilde{\rho}} & \widetilde{\mathcal{H}} \otimes \widetilde{\mathcal{V}} \\ \phi \downarrow & & \downarrow \phi \otimes \tilde{\eta} \\ \mathcal{V} & \xleftarrow{[\cdot]} & \mathcal{H} \end{array}$$

4.5 The substitution law

Given a B-series $B_{hf}(\alpha, y)$, it is of interest to study how this B-series is transformed if hf is substituted by $h\tilde{f}(y) = B_{hf}(\beta, y)$, where $\beta \in \mathfrak{g}$. As shown in [9], [10],

$$B_{h\tilde{f}}(\alpha, y) = B_{hf}(\beta \star \alpha, y),$$

where the coefficients $(\beta \star \alpha)_u$ for $u \in \mathcal{T}$ are given as polynomials in β_v and α_w.

More generally, one may be interested in studying the series of differential operators $S_{h\tilde{f}}(\alpha)$ $(\alpha \in \mathcal{H}^*)$ under the substitution $h\tilde{f}(y) = B_{hf}(\beta, y)$.

Proposition 4.8 *Given $\beta \in \mathfrak{g}$ and $h\tilde{f} = B_{hf}(\beta, y)$, for each $\alpha \in \mathcal{H}^*$,*

$$B_{h\tilde{f}}(\alpha, y) = B_{hf}(\beta \star \alpha, y), \quad S_{h\tilde{f}}(\alpha) = S_{hf}(\beta \star \alpha),$$

where $\beta \star \alpha \in \mathcal{H}^*$ is determined as follows: $\forall u \in \mathcal{H}$,

$$\langle \beta \star \alpha, u \rangle = \langle \alpha, \phi(u) \rangle,$$

where $\phi : \mathcal{H} \to \mathcal{H}$ is the algebra map given by Proposition 4.6 with $\widetilde{\mathcal{H}} = \mathcal{H}$, $\widetilde{\mathcal{V}} = \mathcal{U}$, and $\tilde{\eta}(u) = \beta_u$ for each $u \in \mathcal{T}$.

Hence, $\phi : \mathcal{H} \to \mathcal{H}$ is recursively determined with the help of Proposition 4.7 as follows: For each $u \in \mathcal{T}$,

$$\phi(u) = \beta_u\, [e] + \sum_{\substack{(V \succ W) \subset U \\ V, W \neq \emptyset}} \beta_w\, [\phi(v)].$$

4.6 Bibliographical comments

The commutative Hopf algebra of rooted trees $\mathcal{H}$ was first identified by Dür [14], who realized that Butcher's group [2] was an affine group scheme, and as such, was equivalent to a commutative Hopf algebra. In [16], Grossman and Larson constructed a commutative Hopf algebra $\mathcal{H}_{GL}$ of rooted trees related to the linear differential operators in Subsection 3.3, which was eventually shown [15, 18] to be the graded dual of $\mathcal{H}$ (that is the universal enveloping algebra of $\mathfrak{g}_0$, hence a Hopf subalgebra of the dual Hopf algebra $\mathcal{H}^\circ$). Later on, Connes and Kreimer rediscovered the commutative Hopf algebra $\mathcal{H}$ in the context of renormalization in quantum field theory [22, 13]. Brouder [1] seems to be the first author to note the relationship of Kreimer's work with Butcher's theory. The substitution law in Subsection 4.5 was first considered (in the context of numerical integrators expanded as B-series) in [9] (see also [10]). A bialgebra and a Hopf algebra associated to the substitution law $\star$ (with coproduct obtained by dualizating $\star$) is studied in [6]. Proposition 4.6 can be obtained by combining Chapoton's result [8] (that says that the pre-Lie algebra on the linear span of rooted trees is the free pre-algebra on one generator) and Loday and Ronco's Theorem 5.8 in [23]. Proposition 4.7 is to our knowledge new. Proposition 4.8 can be seen as a dualization of a recursion for the dual of $\star$ found in [6, **?**].

References

[1] C. Brouder, *Runge-Kutta methods and renormalization*, Euro. Phys. J. C 12 (2000), 512–534.

[2] J. C. Butcher, *An algebraic theory of integration methods*, Math. Comput. 26 (1972), 79–106.

[3] J. C. Butcher, *Numerical Methods for Ordinary Differential Equations*, Second edition, J. Wiley, Chichester, (2008).

[4] J. C. Butcher, T. M. H. Chan, *A new approach to the algebraic structures for integration methods*, BIT 42, No. 3 (2002), 477–489.

[5] M. P. Calvo, A. Murua, and J. M. Sanz-Serna, *Modifed equations for ODEs*, Contemp. Math. 172 (1994), 63–74.

[6] D. Calaque, K. Ebrahimi-Fard, D. Manchon, *Two interacting Hopf algebras of trees*, Adv. Appl. Math. (to appear).

[7] T. M. H. Chan, *Algebraic structures for the analysis of numerical methods*, PhD thesis, University of Auckland (1998).

[8] F. Chapoton, M. Livernet, *Pre-Lie algebras and the rooted trees operad*, International Math. Research Notices 8 (2001), 395–408.

[9] P. Chartier, E. Hairer, and G. Vilmart, *A substitution law for B-series vector fields*, Technical Report 5498, INRIA, France, 2005.

[10] P. Chartier, E. Hairer, and G. Vilmart, *Algebraic structures of B-series*, Found. Comput. Math. 10(7) (2010), 407-427.

[11] P. Chartier, E. Faou, A. Murua, *An algebraic approach to invariant preserving integrators: The case of quadratic and Hamiltonian invariants*, Numerische Mathematik, vol. 103 (2006), 575–590.

[12] P. Chartier, A. Murua, *Preserving first integrals and volume forms of additively split systems*, IMA Journal of Numerical Analysis, vol. 27, 3 (2007), 381–405.

[13] A. Connes, D. Kreimer, *Hopf algebras, renormalization, and noncommutative geometry*, Commun. Math. Phys. 199, (1998), 203–242.

[14] A. Dür, *Mobius Functions, Incidence Algebras and Power-Series Representations*, Lecture Notes in Mathematics, 1202, Springer-Verlag, Berlin/Heidelberg, 1986.

[15] L. Foissy, *Les algèbres de Hopf des arbres enracinés décorés, II*, Bulletin des Sciences Mathematiques, 126, 4 (2002), 249–288.

[16] R. Grossman, R. G. Larson, *Hopf-algebraic structure of families of trees*, J. Algebra 126 (1989), 184–210.

[17] R. Grossman, R. G. Larson, *Hopf-algebraic structure of combinatorial objects and differential operators*, Israel Journal of Mathematics 72 (1990), 109–117.

[18] M. E. Hoffman, *Combinatorics of rooted trees and Hopf algebras*, Transactions of the American Mathematical Society 355 (2003), 3795-3811.

[19] E. Hairer, G. Wanner, *On the Butcher group and general multi-value methods*, Computing 13 (1974), 1–15.

[20] E. Hairer, *Backward error analysis of numerical integrators and symplectic methods*, Annals of Numerical Mathematics 1 (1994), 107–132.

[21] G. P. Hochschild, *Basic theory of algebraic groups and Lie algebras*, GTM 75, Springer Verlag, 1981.

[22] D. Kreimer, *On the Hopf algebra structure of perturbative quantum field theories*, Adv. Theor. Math. Phys. 2 (1998), 303–334.

[23] J.-L. Loday and M. Ronco, *Combinatorial Hopf Algebras*, Quanta of maths, Clay Math. Proc., 11, Amer. Math. Soc., Providence, RI, (2010), 347–383.

[24] D. Manchon, *ALGEBRAIC BACKGROUND FOR NUMERICAL METHODS, CONTROL THEORY AND RENORMALIZATION*

[25] D. Manchon and A. Sadi, *Lois Pr-Lie en Interaction*, Communications in Algebra, 39:10 (2011), 3662–3680.

[26] R. H. Merson, *An operational method for the study of integration processes*, Proc. Symp. Data Processing, Weapons Research Establishment, Salisbury, Australia (1957), 110-1 to 110-25.

[27] A. Murua, *Formal series and numerical integrators. Part I: Systems of ODEs and Symplectic integrators*, Appl. Num. Math. 29 (1999), 221–251.

[28] A. Murua, *The Hopf algebra of rooted trees, free Lie algebras, and Lie series*, Found. Comp. Math. 6 (2006), 387–426.

[29] M. E. Sweedler, *Hopf algebras*, W. A. Benjamin, Inc. 1969.