Lie–Butcher series and geometric numerical integration on manifolds

PhD Thesis

Alexander Lundervold

Department of Mathematics
University of Bergen

June, 2011

Acknowledgements

This dissertation is submitted as a partial fulfillment of the requirements for the degree Doctor of Philosophy (PhD) at the Faculty of Mathematics and Natural Sciences, University of Bergen, Norway.

I would like to thank my advisors, Prof. Hans Z. Munthe-Kaas and Dr. Kurusch Ebrahimi-Fard, for their support and guidance throughout my period as a doctoral candidate. Their breadth of mathematical interests and their extensive network of collaborators has helped broaden my mathematical horizon considerably.

A special thanks also to the students and staff of the Department of Mathematics for making my time there well spent.

Contents

Outline of the thesis

The thesis belongs to the field of "geometric numerical integration" (GNI), whose aim it is to construct and study numerical integration methods for differential equations that preserve some geometric structure of the underlying system. Many systems have conserved quantities, e.g. the energy in a conservative mechanical system or the symplectic structures of Hamiltonian systems, and numerical methods that take this into account are often superior to those constructed with the more classical goal of achieving high order.

An important tool in the study of numerical methods is the *Butcher series* (B-series) invented by John Butcher in the 1960s. These are formal series expansions indexed by rooted trees and have been used extensively for order theory and the study of structure preservation. The thesis puts particular emphasis on B-series and their generalization to methods for equations evolving on manifolds, called *Lie–Butcher series* (LB-series).

It has become apparent that algebra and combinatorics can bring a lot of insight into this study. Many of the methods and concepts are inherently algebraic or combinatoric, and the tools developed in these fields can often be used to great effect. Several examples of this will be discussed throughout.

The thesis is structured as follows: background material on geometric numerical integration is collected in **Part I**. It consists of several chapters: in **Chapter 1** we look at some of the main ideas of geometric numerical integration. The emphasis is put on B-series, and the analysis of these. **Chapter 2** is devoted to differential equations evolving on manifolds, and the series corresponding to B-series in this setting. **Chapter 3** consists of short summaries of the papers included in Part II. **Part II** is the main scientific contribution of the thesis, consisting of reproductions of three papers on material related to geometric numerical integration.

Part I

Background

Chapter 1

Geometric numerical integration on vector spaces

In numerical analysis the main objects of study are flows of vector fields, given by initial value problems of the type*:

$$y'(t) = F(y(t)), \quad y(t_0) = y_0. \tag{1.1}$$

The function y can be real-valued or vector-valued (giving rise to a system of coupled differential equations). The flow of the differential equation is the map $\Psi_{t,F} : \mathbb{R}^n \to \mathbb{R}^n$ defined by $y(t) = \Psi_{t,F}(y_0)$.† Note that $F(y) = d/dt|_{t=0}\Psi_{t,F}(y_0)$. In many practical settings, for instance many mechanical systems modeling physical processes, the vector field is Hamiltonian, and such flows have several interesting geometric properties. We seek to construct good approximations to the exact flow, where 'good' can mean several different things, depending on the context. Sometimes what we want are integrators of high order, other times we need approximations that preserve some qualitative or geometric structure of the underlying dynamical system. Preserving geometric structure is particularly important when studying systems over long time intervals. An early illustration of this fact was made by Wisdom and Holman in [75], where they computed the evolution of the solar system over a billion-year time period using a symplectic method, making an energy-error of only 2×10^{-11}. Section 1.1 of this thesis focuses on structure preservation for numerical methods.

As there are several excellent introductions to geometric numerical integration on $\mathbb{R}^n$ we will not go into a detailed study here, but merely describe some of the main ideas. The book [35] is the standard reference; other introductions can be found in [54, 45, 5, 53, 64, 69, 71].

The focus of this thesis will be on some of the algebraic and combinatorial tools of geometric numerical integration, with particular emphasis on the tools we

* Non-autonomous differential equations can also be written on this form by adding a component to the y vector

† Here we assume Lipschitz continuity of F for the flow to exist and be unique.

will utilize when studying flows on more general manifolds in the next chapter. Lately, there has been quite a lot of interest in these algebraic aspects of geometric integration, and this has resulted in both an increased understanding of the field, and also of its relations to other areas of mathematics.

1.1 Numerical methods and structure-preservation

Consider an initial value problem of the form (1.1):

$$y'(t) = F(y(t)), \quad y(0) = y_0$$

representing the flow of the (sufficiently smooth) vector field F. A numerical method for (1.1) generates approximations $y_1, y_2, y_3, \dots$ to the solution $y(t)$ at various values of t. One of the simplest methods is the (explicit) **Euler method**. It computes approximations y_n to the values $y(nh)$, where $n \in \mathbb{N}$ and h is the step size, using the rule:

$$y_{n+1} = y_n + hF(y_n). \tag{1.2}$$

This generates a numerical flow Φ_h approximating the exact flow Ψ of F. The accuracy of the method can be measured by its **order**: we say that a one-step method $y_{n+1} = \Phi_h(y_n)$ has order n if $|\Phi_h(y) - \Psi_h(y)| = O(h^{n+1})$ as $h \to 0$. Another way to put this is in terms of the curve traced out by the numerical flow: by comparing its Taylor series to the Taylor series for the curve of the exact flow term by term, we can read off the order of the method. The Taylor series for the solution y has the form

$$y(h) = y_0 + hF(y_0) + \frac{1}{2}h^2 F'(y_0)F(y_0) + O(h^3),$$

and we note that the Euler method is of order 1.

Runge–Kutta methods. The Euler method is an example of a **Runge–Kutta method**, a class of methods that are very common in applications [36, 8]. A Runge–Kutta method is a one-step method computing an approximation y_1 to $y(h)$ with y_0 as input, as follows:

Definition 1.1. An s-stage Runge–Kutta method for solving the initial value problem (1.1) is a one-step method given by

$$\begin{aligned} Y_i &= y_0 + h\sum_{j=1}^{s} a_{ij}F(Y_j), \quad i = 1, \dots s \\ y_1 &= y_0 + h\sum_{i=1}^{s} b_i F(Y_i), \end{aligned} \tag{1.3}$$

where $b_i, a_{ij} \in \mathbb{R}$, h is the step size and $s \in \mathbb{N}$ denotes the number of *stages*.

A Runge–Kutta method can be presented as a *Butcher tableau*, which characterizes the method completely:

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array}$$

Here $c_i = \sum_{j=1}^{s} a_{ij}$.

Example 1.2. We note that the Euler method is the Runge–Kutta method with Butcher tableau:

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}$$

Another well-known example is the **explicit midpoint method**:

$$y_{n+1} = y_n + hF\left(y_n + \frac{1}{2}hF(y_n)\right),$$

given by:

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

Given any number m, there is a Runge–Kutta method of order m [8]. Verifying this involves expanding the methods into series involving the derivatives of F, and already at low orders the expressions get quite complicated. However, in Section 1.2 we shall see that the Runge–Kutta methods are special cases of *Butcher series methods*, and that one can find nice descriptions of the order theory and also structure preservation properties for numerical methods within this framework.

Differential equations and geometric structures. When presented with a system modeled by a differential equation one will often first try to determine its qualitative properties: are there any invariants? What kind of geometric structure does the system have? Structures of interest can be energy and volume preservation, symplectic structure, first integrals, restriction to a particular manifold (as studied in Chapter 2), etc. Then, when choosing (or designing) a numerical method for approximating the solution of the differential equation, it might make sense for the method to share these qualitative features. In that way one has control over what kind of errors the method introduces, obtaining a method tailor-made to the problem at hand.

A rich source of problems with geometric structures are the **Hamiltonian systems**. Let $H : \mathbb{R}^{2n} \to \mathbb{R}$ be a smooth function. A **Hamiltonian vector field** is a vector field on $\mathbb{R}^{2n}$ of the form $X_H = \Omega^{-1}\nabla H$, where Ω is an antisymmetric, invertible $2n \times 2n$ matrix.‡ The flow of X_H is given by

$$\frac{d}{dt}z = \Omega\nabla_z H(z).$$

The function H represents the total energy of the system. Two important properties of the flow of a Hamiltonian vector field X_H is that it is constant along the Hamiltonian function H (conservation of energy) and that it preserves a symplectic form ω on R^{2n}. Using numerical integrators constructed to preserve these properties has been shown to lead to dramatic improvements in accuracy. For examples of this phenomenon see e.g. [35, 34, 45] and references therein.

1.2 Trees and Butcher series

Starting with the work of John Butcher in the 1960s and 70s [6, 7] the study of methods for solving ordinary differential equations has been closely connected to the combinatorics of rooted trees. Many numerical methods $y_{n+1} = \Phi_h(y_n)$ (including all Runge–Kutta methods) can be expressed as certain formal series, named **Butcher series** by Hairer and Wanner in [37]. By a clever representation of the terms, the series can be indexed over the set of rooted trees.

Consider the differential equation

$$y'(x) = F(y(x)). \tag{1.4}$$

Denote the components of $F : \mathbf{R}^n \to \mathbf{R}^n$ by f^i and write

$$f^i_{j_1 j_2 \cdots j_k} = \frac{\partial^k f^i}{\partial x_{j_1}\partial x_{j_2}\cdots\partial x_{j_k}}. \tag{1.5}$$

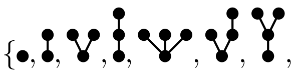
Summing over repeated indices, the first few derivatives of y can be written as:

$$\begin{aligned}
\frac{dy^i}{dx} &= f^i \\
\frac{d^2y^i}{dx^2} &= f^i_j f^j \\
\frac{d^3y^i}{dx^3} &= f^i_{jk} f^j f^k + f^i_j f^j_k f^k \\
\frac{d^4y^i}{dx^4} &= f^i_j f^j_k f^k_l f^l + f^i_j f^j_{kl} f^k f^l + 3 f^i_{jk} f^j_l f^k f^l + f^i_{jkl} f^j f^k f^l.
\end{aligned} \tag{1.6}$$

These expressions soon get very complicated, but the structure can be made much more transparent by observing that the derivatives of F can be associated in a bijective way with rooted trees, an observation already made by Cayley in 1857 [14]. Before giving the exact correspondence between differential equations, rooted trees and Butcher series, we will take a closer look at trees.

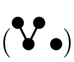
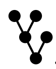
‡ Hamiltonian vector fields can be defined on any symplectic manifold [3].

Rooted trees. A **tree** is a connected graph with no cycles

$T = \{$  $, \ldots\}.$

A **rooted tree** is a tree with one vertex designated as the root. In the pictorial representation of trees, the root will always be drawn as the bottom vertex, and the trees will be ordered from the root to the top. More precisely, a tree τ is a graph consisting of a set of vertices $V(\tau)$ and edges $E(\tau) \subset V(\tau) \times V(\tau)$ so that there is exactly one path connecting any two vertices. A **path** between v_i and v_j is a set of edges $\{v_{s_l}, v_{t_l}\}$ so that $l = 1, 2, \ldots, r$, $s_1 = i$, $t_l = s_{l+1}$ and $t_r = j$. This gives a partial ordering of the tree in terms of paths from the root to the vertices of the tree. A vertex v_i is smaller than another distinct vertex v_j, e.g. $v_i \prec v_j$, if the unique path from from the root to v_j goes via v_i. A vertex v_i is called a **leaf** if there is no vertex v_j with $v_i \prec v_j$. A **child** of a vertex v_i is a vertex v_j with $v_i \prec v_j$ so that there is no vertex v_k with $v_i \prec v_k \prec v_j$. The **order** $|\tau|$ of a tree τ is the number of vertices of the tree. We define a symmetry group on a tree τ as all automorphisms on the vertices. The order of this group, $\sigma(\tau)$, is called the **symmetry** of the tree τ.

A **forest** of rooted trees is a graph whose connected components are rooted trees, e.g. $\omega = \tau_1 \ldots \tau_n$. We include the *empty tree* $\mathbb{I}$, i.e. the graph with no vertices, in the set F of forests. F can be put in bijection to the set of trees via the operator $B^+ : F \to T$, defined on a forest $\omega = \tau_1 \ldots \tau_n$ by connecting the trees to a new root by addition of edges. For example,

$B^+($  $) =$  .

This operator can be used to generate all trees recursively from the tree $\bullet$ by the following procedure:

(i) The graph $\bullet$ belongs to T

(ii) If $\tau_1, \ldots, \tau_n \in$ T then $\tau = B^+(\tau_1 \ldots \tau_n)$ is in T.

The **tree factorial** $\tau!$ is given recursively by:

(i) $\bullet! = 1$

(ii) $B^+(\tau_1 \ldots \tau_n)! = |B^+(\tau_1 \ldots \tau_n)|\tau_1! \ldots \tau_n!$.

An important operation on trees is the Butcher product, defined in terms of *grafting*.

Definition 1.3. The **Butcher product** $\tau \diamond \omega$ of a tree $\tau = \mathrm{B}^+(\tau_1 \ldots \tau_n)$ and a forest $\omega = \omega_1 \cdots \omega_m$ is given by grafting ω onto the root of τ:

$$\tau \diamond \omega = \mathrm{B}^+(\tau_1 \ldots \tau_n \, \omega_1 \ldots \omega_m) \tag{1.7}$$

Butcher series. The calculations of the derivatives of $y'(t) = F(y(t))$ performed at the beginning of the section can be written in terms of the elementary differentials of F.

Definition 1.4. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field. The **elementary differential** $\mathcal{F}$ of F is

$$\begin{aligned}\mathcal{F}(\bullet)(t) &= F(y)\\ \mathcal{F}(\tau)(t) &= F^{(m)}(y)(\mathcal{F}(\tau_1)(y), \dots, \mathcal{F}(\tau_m)(y)),\end{aligned} \tag{1.8}$$

where $F^{(m)}$ is the m-th derivative of the vector field F and $\tau = B^+(\tau_1, \dots, \tau_m)$ is a rooted tree.

We will discuss another way to write elementary differentials in Section 1.5. With the notation from Equation (1.5), the first few elementary differentials are shown in Table (1.1). The vector field F corresponds to the leaves of the tree, the first derivative F' corresponds to a vertex with an edge with one child, the second derivative F'' corresponds to a vertex with two children, etc.

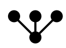
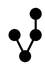
τ	$\mathcal{F}(\tau)(y)^i$
$\bullet$	f^i
	$f^i_j f^j$
	$f^i_{jk} f^j f^k$
	$f^i_j f^j_k f^k$
	$f^i_{jkl} f^j f^k f^l$
	$f^i_{jk} f^j f^k_l f^l$

Table 1.1: Elementary differentials associated to a vector field F with components f^i.

Butcher series are (formal) Taylor expansions of elementary differentials indexed over trees:

Definition 1.5. A **Butcher series** (B-series) is a (formal) series expansion in a parameter h:

$$\begin{aligned}\mathcal{B}_{h,F}(\alpha) &= \alpha(\mathbb{I})\mathcal{F}(\mathbb{I}) + \sum_{\tau \in \mathrm{T}} h^{|\tau|} \frac{\alpha(\tau)}{\sigma(\tau)} \mathcal{F}(\tau)\\ &= \sum_{\tau \in \tilde{\mathrm{T}}} h^{|\tau|} \frac{\alpha(\tau)}{\sigma(\tau)} \mathcal{F}(\tau),\end{aligned} \tag{1.9}$$

where $\tilde{\mathrm{T}} = \mathrm{T} \cup \{\mathbb{I}\}$, F is a vector field, α is a function $\alpha : \tilde{\mathrm{T}} \to \mathbb{R}$, $\sigma(\tau)$ is the symmetry of τ, h is a real number (representing the step size), and $\mathcal{F}$ is the elementary differential of F, extended to the empty tree $\mathbb{I}$ by $\mathcal{F}(\mathbb{I})(y) = y$.

We shall see that these series can be used to represent numerical methods $y_{n+1} = \Phi_h(y_n)$ approximating the flow of a vector field F, in the sense that the Taylor series for Φ_h can be expanded into a B-series: $\Phi_h = \mathcal{B}_{h,F}(\alpha)$. §

By computing the Taylor expansion of the solution to the initial value problem (1.1) one obtains the following result:

Proposition 1.6 ([35])**.** The Taylor series for the solution of the differential equation (1.1) can be written as a B-series:

$$B_{h,F}(\gamma) = \sum_{\tau \in \tilde{\mathrm{T}}} h^{|\tau|} \frac{\gamma(\tau)}{\sigma(\tau)} \mathcal{F}(\tau), \tag{1.10}$$

where $\gamma(\tau) = 1/\tau!$. That is, $y(t+h) = \mathcal{B}_{h,F}(\gamma)(y(t))$.

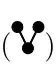
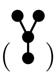
Runge–Kutta methods can also be written as B-series expansions, with coefficients given by the *elementary weights* of the method [6].

Definition 1.7 (Elementary weights)**.** Let b_i and a_{ij} be coefficients of a RK-method as in Definition 1.1, where $i \in \mathbb{N}$. The **elementary weight function** Φ is defined on trees as follows:

$$\begin{aligned}
\Phi_i(\bullet) &= c_i \\
\Phi(\bullet) &= \sum_{j=1}^{s} b_j \\
\Phi_i(B^+(\tau_1, \ldots, \tau_k)) &= \sum_{j=1}^{s} a_{ij} \Phi_j(\tau_1)\Phi_j(\tau_2)\ldots\Phi_j(\tau_k) \\
\Phi(B^+(\tau_1, \ldots, \tau_k)) &= \sum_{j=1}^{s} b_j \Phi_j(\tau_1)\Phi_j(\tau_2)\ldots\Phi_j(\tau_k)
\end{aligned} \tag{1.11}$$

Here $i = 1, \ldots, s$.

For example,

$\Phi($  $) = \sum_{j=1}^{s} b_j c_j,\quad \Phi($  $) = \sum_{j=1}^{s} b_j c_j^2,\quad \Phi($  $) = \sum_{j,k=1}^{s} b_j a_{jk} c_k^2$

Theorem 1.8 ([6])**.** *The B-series for a RK-method given by the elementary weights* $\Phi(\tau)$ *is*

$$\mathcal{B}_{h,F}(\Phi) = \sum_{\tau \in \tilde{\mathrm{T}}} h^{|\tau|} \frac{\Phi(\tau)}{\sigma(\tau)} \mathcal{F}(\tau) \tag{1.12}$$

§ A numerical method for solving a differential equation is called a *B-series method* if it can be written as a B-series.

Order theory for B-series methods. Once we have the B-series of the exact solution and the B-series of a numerical method, it is straightforward to compare the coefficients and read off the order of the method. For Runge–Kutta methods, we obtain the following result:

Proposition 1.9 ([6])**.** A Runge–Kutta method given by a B-series with coefficients $\Phi(\tau)$ has order n if and only if

$$\Phi(\tau) = \gamma(\tau), \qquad \text{for all } \tau \in T \text{ such that } |\tau| < n.$$

B-series methods and structure preservation. The class of B-series methods includes all Taylor series methods and Runge–Kutta methods. It does not, however, include all numerical methods, an example being the class of *splitting methods*.

It is important to point out that focusing only on B-series methods has its drawbacks. Besides the fact that the class does not contain all methods, it is also known that there are certain geometric structures that cannot be preserved by B-series methods. For example, no B-series method can preserve the volume for *all* systems [41]. However, we will be content with this loss of generality and focus exclusively on methods based on B-series in this chapter, and on their generalization – Lie–Butcher series – in the next.

A case which is particularly well-studied is Hamiltonian vector fields. The following two theorems serve as prime examples:

Theorem 1.10 ([33])**.** *Let $G = \mathcal{B}_{h,F}(\alpha)$ be a vector field with $\alpha(\mathbb{I}) = 0$, $\alpha(\bullet) \neq 0$. Then G is Hamiltonian for all Hamiltonian vector fields $F(y) = \Omega^{-1}\nabla H(y)$ if and only if*

$$\alpha(\tau_1 \diamond \tau_2) + \alpha(\tau_2 \diamond \tau_1) = 0 \tag{1.13}$$

for all $\tau_1, \tau_2 \in \mathrm{T}$. Here $\diamond$ denotes the Butcher product of Definition 1.3.

Theorem 1.11 ([12])**.** *Consider a numerical method given by a B-series $\mathcal{B}_{h,F}(\alpha)$. The method is symplectic if and only if*

$$\alpha(\tau_1 \diamond \tau_2) + \alpha(\tau_2 \diamond \tau_1) = \alpha(\tau_1)\alpha(\tau_2) \tag{1.14}$$

for all $\tau_1, \tau_2 \in \mathrm{T}$, where $\alpha(\mathbb{I}) = 0$.

The paper [16] gives an overview of what is known about structure preservation for B-series, including characterizations of the various subsets of trees corresponding to energy-preserving, Hamiltonian and symplectic B-series.

1.3 Hopf algebras and the composition of Butcher series

Consider two numerical methods given by Φ^1 and Φ^2. Using the method Φ^1 to advance a point y_0 to a point y_1, and then applying the method Φ^2 using y_1 as initial point, results in a point y_2:

$$y_1 = \Phi^1(y_0), \qquad y_2 = \Phi^2(y_1).$$

This is the idea behind **composition** of numerical methods. In the case where both methods are given by B-series, $\Phi^1(y_1) = \mathcal{B}^1_{h,F}(\alpha)(y_0)$, $\Phi^2(\tilde{y}_1) = \mathcal{B}^2_{h,F}(\beta)(\tilde{y}_0)$, the composition method $\Phi^2 \circ \Phi^1$ is again a B-series: $\Phi^2 \circ \Phi^1(y_0) = \mathcal{B}_{h,F}(\gamma)(y_0)$. This is the Hairer–Wanner theorem from [37]. The coefficient function γ of this B-series was first studied by John Butcher in [7], where he found that composition of B-series is a group operation (giving rise to the *Butcher group*) on the coefficient functions, and gave expressions for the product, identity and inverse in this group.

In [43, 21] Connes and Kreimer introduced a Hopf algebra of rooted trees connected to the renormalization procedure in quantum field theory. Later [4] it was pointed out that a variant of this Hopf algebra is closely related to the Butcher group. More precisely, the Butcher group is the group of *characters* in a Hopf algebra $\mathrm{H}_{\mathrm{BCK}}$ defined by Connes and Kreimer.

We will describe the Butcher group indirectly by describing the Hopf algebra $\mathrm{H}_{\mathrm{BCK}}$. But first we will present some basic definitions from the theory of Hopf algebras. For a comprehensive introduction, see [68, 1]. Other excellent references include [13, 51]. A short introduction can also be found in Paper A, reprinted in Part II below.

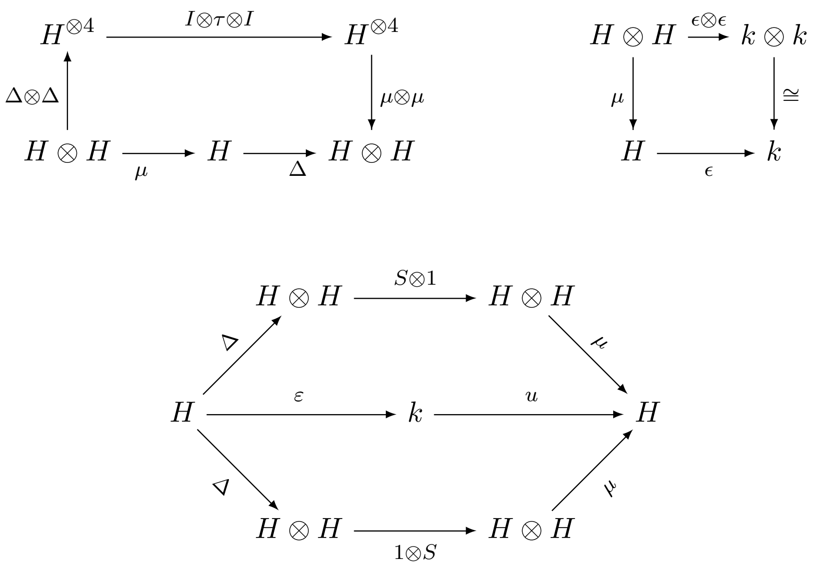
Hopf algebras. Let k be a field of characteristic zero. An **algebra** A over k is a k-vector space equipped with a multiplication map $\mu : A \otimes A \to A$ and a unit $u : \mathrm{k} \to A$ so that

- $\mu \circ (id \otimes \mu) = \mu \circ (\mu \otimes id) : A \otimes A \otimes A \to A$ (associativity)
- $\mu \circ (u \otimes id) = \mu \circ (id \otimes u) : k \otimes A \cong A \to A$ (unitality)

A **coalgebra** C over k is the dual notion. It consists of a comultiplication map $\Delta : C \to C \otimes C$ and a counit $\epsilon : C \to \mathrm{k}$ so that

- $(\Delta \otimes id) \circ \Delta = (id \otimes \Delta) \circ \Delta : C \to C \otimes C \otimes C$ (coassociativity)
- $(\epsilon \otimes id) \circ \Delta = (id \otimes \epsilon) \circ \Delta : C \to C \otimes \mathrm{k} \cong C$ (counitality)

A **Hopf algebra** is at once an algebra and a coalgebra, and it comes equipped with an antipode $S : H \to H$. These structures have to satisfy certain compatibility conditions, written as the following diagrams, where τ denotes the flip operation $\tau(h_1, h_2) = (h_2, h_1)$:



The first two diagrams ensure that the coproduct and the counit are both algebra homomorphisms. The last diagram is best interpreted in terms of the characters in a Hopf algebra. Let A be a commutative k-algebra, and let $\mathcal{L}(H,A)$ denote the set of linear maps from H to A. An element $\alpha \in \mathcal{L}(H,A)$ is called a **character** if $\alpha(x\cdot y) = \alpha(x)\cdot\alpha(y)$ for all $x, y \in H$, where the product on the left-hand side is in H, and on the right-hand side in A. The set of characters in $\mathcal{L}(H,A)$ form a group under the **convolution product**:

$$\phi * \psi = \mu \circ (\phi \otimes \psi) \circ \Delta. \tag{1.15}$$

The unit is the composition of the unit and the counit in H, e.g. $\eta := u \circ \epsilon$. The bottom diagram above corresponds to the antipode being the inverse of the identity under this product, and we have $\alpha^{*-1} = \alpha \circ S$.

We will also need the concept of **infinitesimal characters**, which are maps α in $\mathcal{L}(H,A)$ satisfying

$$\alpha(x\cdot y) = \eta(x)\cdot\alpha(y) + \alpha(x)\cdot\eta(y).$$

The Butcher–Connes–Kreimer Hopf algebra. Composition of B-series is governed by a certain Hopf algebra $\mathrm{H_{BCK}}$ based on the set T of rooted trees, called the *Butcher-Connes-Kreimer Hopf algebra*. In the next chapter we will see that a generalization of this Hopf algebra governs the composition of Lie-Butcher series (Section 2.2.3).

To describe the BCK Hopf algebra we need to define its structure as a vector space, an algebra, a coalgebra, and define the antipode. As a $\mathbb{R}$-vector space $\mathrm{H_{BCK}}$ is generated by the set T of rooted trees, and graded by the order (i.e. number of vertices) of the trees. The algebra structure is that of the symmetric algebra

$S(\mathbb{R}\{T\})$. The product is written as (commutative) concatenation of trees (i.e. disjoint union), giving rise to forests of trees. The unit is the empty tree $\mathbb{I}$.

The coproduct of $\mathrm{H}_{\mathrm{BCK}}$ is the map $\Delta_{\mathrm{BCK}} : \mathrm{H}_{\mathrm{BCK}} \to \mathrm{H}_{\mathrm{BCK}} \otimes \mathrm{H}_{\mathrm{BCK}}$ determined recursively by:

$$\Delta_{\mathrm{BCK}} \circ B^{+}(\omega) = B^{+}(\omega) \otimes \mathbb{I} + (Id \otimes B^{+}) \circ \Delta_{\mathrm{BCK}}(\omega), \tag{1.16}$$

where ω is a forest¶. The counit is the map $\epsilon : \mathrm{H}_{\mathrm{BCK}} \to \mathbb{R}$ given by $\epsilon(\mathbb{I}) = 1$ and $\epsilon(\tau) = 0$ if $\tau \neq \mathbb{I}$. The coproduct can also be written in a non-recursive manner using cuttings of trees.

Cutting trees. An **admissible cut** of a tree τ is a set $c \subset E(\tau)$ of edges of τ such that c contains at most one edge from any path from the root to a leaf. The case $c = \emptyset$ is called the empty cut. The cut $c = E(\tau)$ is allowed, and is called the full cut. Let ω denote the forest with vertices $V(\tau)$ and edges $E(\tau) \setminus c$. We write $R^{c}(\tau)$ for the component of ω containing the root of τ, and $P^{c}(\tau)$ for the forest consisting of the remaining components.

Theorem 1.12 ([21]). *The coproduct in* $\mathrm{H}_{\mathrm{BCK}}$ *can be written as*

$$\Delta_{\mathrm{BCK}}(\tau) = \sum_{c \in Adm(\tau)} P^{c}(\tau) \otimes R^{c}(\tau) \tag{1.17}$$

Examples of the coproduct can be found in Table 1.2. The antipode can be defined recursively as:

$$S(\tau) = -\tau - \sum_{c \in Adm(\tau) \setminus \emptyset} S(P^{c}(\tau)) R^{c}(\tau) \tag{1.18}$$

The Hairer–Wanner theorem gives the exact correspondence between $\mathrm{H}_{\mathrm{BCK}}$ and composition of B-series:

Theorem 1.13 ([37]). *Let* $\mathcal{B}^{1}_{h,F}(\alpha)$ *and* $\mathcal{B}^{2}_{F}(\beta)$ *be two B-series, with coefficients* $\alpha, \beta : T \to \mathbb{R}$. *The composition* $\mathcal{B}^{2}_{h,F}(\beta) \circ \mathcal{B}^{1}_{h,F}(\alpha)$ *is again a B-series, and we have*

$$\mathcal{B}^{2}_{h,F}(\beta) \circ \mathcal{B}^{1}_{h,F}(\alpha) = \mathcal{B}_{h,F}(\alpha \star \beta), \tag{1.19}$$

where $\star$ *denotes convolution in the Hopf algebra* $\mathrm{H}_{\mathrm{BCK}}$.

¶ Recall that Δ_{BCK} is an algebra morphism and is therefore defined on forests as well as trees, since $\Delta_{\mathrm{BCK}}(\tau_1\tau_2) = \Delta_{\mathrm{BCK}}(\tau_1)\Delta_{\mathrm{BCK}}(\tau_2)$.

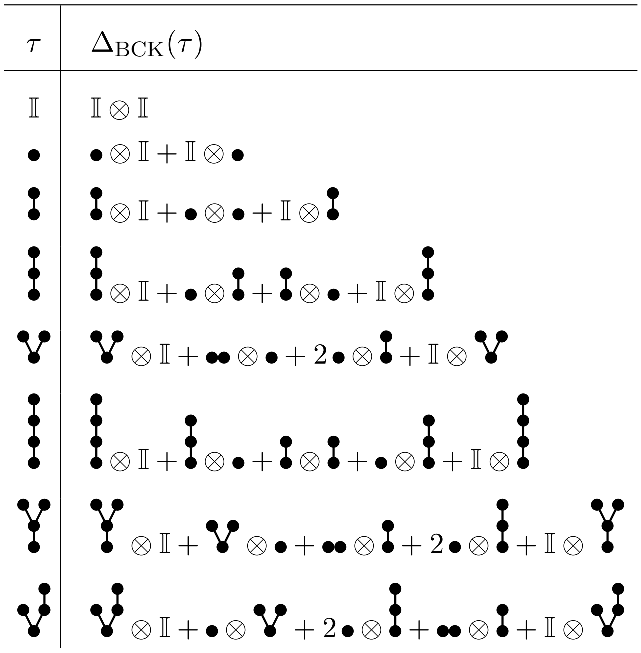


Table 1.2: Examples of the coproduct Δ_{BCK} in the Hopf algebra $\mathrm{H}_{\mathrm{BCK}}$

1.4 Substitution and backward error analysis for Butcher series

Consider a numerical method Φ_h used to solve a differential equation of the form

$$y' = F(y). \tag{1.20}$$

The basic idea of **backward error analysis** of the method Φ_h is to interpret it as giving the exact solution of a modified equation:

$$\tilde{y}' = \tilde{F}_h(\tilde{y}). \tag{1.21}$$

If we can find such an equation, we can use it to study the properties of the numerical method. In other words, the numerical method Φ_h will be represented by a modified vector field $\tilde{F}$, which then can be used to study the method. The idea is based on work by Wilkinson in the context of algorithms for solving equations given by matrices [74], and has been explored in several papers [73, 33, 11, 35, 20]. Recurrence formulas for the modified equation was first obtained in [33, 11].

A related notion is the **modifying integrators** of [20]. The idea is to look for a vector field $\tilde{F}_h$ so that the numerical method Φ_h applied to the flow equation of $\tilde{F}_h$ (Equation 1.21) is the exact solution of Equation 1.20.

It turns out that the case where Φ_h is a B-series method is particularly nice [19, 20, 9]. The vector fields $\tilde{F}_h$ can then be written as B-series whose coefficients are derived from the coefficients of Φ_h, and these coefficients can be expressed by the **substitution law** for B-series methods.

The substitution law. Let $\mathcal{B}_{h,F}(\alpha)$ and $\mathcal{B}_{h,G}(\beta)$ be two B-series, where $\alpha(\mathbb{I}) = 0$. Then $\mathcal{B}_{h,F}(\alpha)$ is a vector field, and we can consider the B-series obtained by using this as the vector field G in the B-series $\mathcal{B}_{h,G}(\beta)$. This is called *substitution* of B-series. The result is given in terms of a bialgebra $\mathrm{H}_{\mathrm{CEFM}}$ by the following theorem:

Theorem 1.14 ([9]). *Let F be a vector field, α, β linear maps $\alpha, \beta : \mathrm{T} \to \mathbb{R}$ where β is an infinitesimal character of $\mathrm{H}_{\mathrm{BCK}}$, and $\alpha(\mathbb{I}) = 0$. Then the vector field $(1/h)\mathcal{B}_{h,F}(\alpha)$ inserted into the B-series $\mathcal{B}_{h,\cdot}(\beta)$ is again a B-series, given by*

$$\mathcal{B}_{h,(1/h)\mathcal{B}_{h,F}(\alpha)}(\beta) = \mathcal{B}_{h,F}(\alpha * \beta), \tag{1.22}$$

where $*$ denotes convolution of characters in the bialgebra H_{CEFM}.

The bialgebra $\mathrm{H}_{\mathrm{CEFM}}$ is the symmetric algebra over rooted trees $S(\mathrm{T})$, with $\bullet$ as unit, equipped with a coproduct given by contracting subforests in trees:

$$\Delta(\tau) = \sum_{\omega \subseteq \tau} \omega \otimes \tau/\omega. \tag{1.23}$$

If τ is a tree then the notation $\omega \subset \tau$ means that ω is a spanning subforest of τ, i.e. that ω is a collection of subtrees of τ so that each vertex of τ belongs to exactly one tree in ω. Then τ/ω denotes the tree obtained by contracting each subtree (with at least two vertices) of τ contained in ω onto a vertex. Some examples of the coproduct can be found in Table 1.3.

There is a Hopf algebra related to $\mathrm{H}_{\mathrm{CEFM}}$, obtained by considering the symmetric algebra over the set of rooted trees T' with at least one edge (e.g. $\bullet$ is not included), and then adding $\bullet$ back as the *unit* for the product. The coproduct is defined as in Equation (1.23). This makes the associated bialgebra connected, and it is therefore a Hopf algebra [51].

For details on these constructions, consult [9].

Backward error analysis and modifying integrators. Once Theorem 1.14 is established one can obtain expressions for backward error analysis and modifying integrators.

Corollary 1.15 (Backward error analysis). *Let $\mathcal{B}_G(\gamma)$ denote the B-series for the exact flow of the vector field G, and let $\mathcal{B}_F(\alpha)$ be a B-series giving a numerical flow for F. The modified vector field $\tilde{F}$ given by $\mathcal{B}_{\tilde{F}}(\gamma) = \mathcal{B}_F(\alpha)$ is a B-series $\mathcal{B}_F(\beta)$ with coefficients given by*

$$\beta * \gamma = \alpha$$

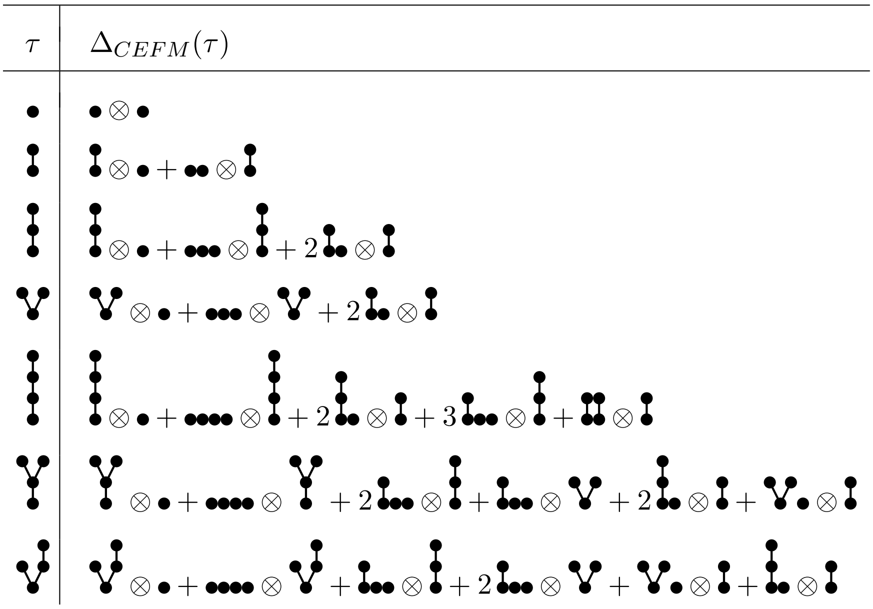


Table 1.3: Examples of the coproduct Δ_{CEFM} in the substitution bialgebra

Corollary 1.16 (Modifying integrators). *Let $\mathcal{B}_G(\gamma)$ denote the B-series for the exact flow of the vector field G, and let $\mathcal{B}_F(\alpha)$ be a B-series giving a numerical flow for F. The modified vector field $\tilde{F}$ so that $\mathcal{B}_{\tilde{F}}(\alpha) = \mathcal{B}_F(\gamma)$ is a B-series $\mathcal{B}_F(\beta)$ whose coefficients are given by*

$$\beta * \alpha = \gamma$$

1.5 Pre-Lie Butcher series

The space of vector fields has the structure of a **pre-Lie algebra**, and in this section we will see that B-series can be formulated purely in terms of this pre-Lie structure. This allows us to lift the concept of B-series to the free pre-Lie algebra, giving rise to **pre-Lie B-series** [26]. Viewing B-series as objects in the free pre-Lie algebra gives a clearer focus on the core algebraic structures at play, and it also enables the application of tools and results from other fields where pre-Lie algebras appear. Two examples of this phenomenon can be found in [25] (see Remark 1.23) and [9]. We give the basic constructions here because formulating Butcher series in terms of pre-Lie algebras will find an analogue in the next chapter, where Lie–Butcher series will be constructed from the so-called D-algebras.

Pre-Lie algebras. The concept of pre-Lie algebras is a relaxation of associative algebras that still preserve their *Lie admissible* property. In other words, for an associative algebra $(A,*)$ antisymmetrization of the product $*$ gives a Lie bracket, making it a Lie algebra: $[a,b]=a*b-b*a$, and this property also holds for pre-Lie algebras. Note, however, that not all pre-Lie algebras are associative. They were first introduced and studied by Vinberg [72], Gerstenhaber [31], and Agrachev and Gamkrelidze [2], under various names. A nice introduction to pre-Lie algebras can be found in [52].

Definition 1.17. A (left) **pre-Lie algebra**[‖] $(A,\triangleright)$ is a k-vector space A equipped with an operation $\triangleright : A\otimes A\to A$ subject to the following relation:

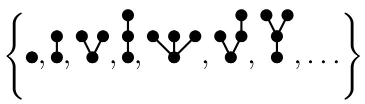
$$(a\triangleright b)\triangleright c-a\triangleright(b\triangleright c) \quad=\quad (b\triangleright a)\triangleright c-b\triangleright(a\triangleright c) \tag{1.24}$$

Example 1.18 (The pre-Lie algebra of vector fields)**.** The space of vector fields $\mathcal{X}(M)$ on a differentiable manifold M equipped with a flat, torsion-free connection ∇ can be given the structure of a pre-Lie algebra by defining $\triangleright$ as $F\triangleright G=\nabla_F G$. In the case $M=\mathbb{R}^n$ with the standard flat and torsion-free connection we have that for $F=\sum_{i=1}^n F_i\partial_i$ and $G=\sum_{j=1}^n G_j\partial_j$,

$$F\triangleright G=\sum_{i=1}^n\left(\sum_{j=1}^n F_j(\partial_j G_i)\right)\partial_i. \tag{1.25}$$

In the next chapter we will see that allowing for torsion leads to the concept of **D-algebras**. See also [48], included as Paper C in Part II of the thesis.

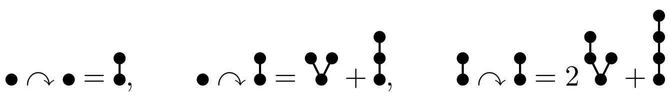
The free pre-Lie algebra. The free pre-Lie algebra has been studied in several papers, most notably by Chapoton and Livernet in [18], Segal in [65], Agrachev and Gramkrelidze in [2], Dzhumadil'daev and Löfwall in [23]. These papers give different bases for the free pre-Lie algebra, and one can choose to work in the basis most beneficial for the problem at hand. A basis for the free pre-Lie algebra $PL(V)$ over a vector space V was described by Chapoton and Livernet in terms of nonplanar rooted trees [18, 17]:



decorated by elements of V. The pre-Lie product $\tau_1\curvearrowright\tau_2$ of two rooted trees is given by grafting: $\tau_1\curvearrowright\tau_2$ is the sum of all the trees resulting from the addition of an edge from the root of τ_1 to one of the vertices of τ_2:

$$\tau_1\curvearrowright\tau_2:=\sum_{v\in V(\tau_2)}\tau_1\circ_v\tau_2 \tag{1.26}$$

Here $\tau_1\circ_v\tau_2$ denotes grafting at the vertex v of τ_2.



‖ Also called a *Vinberg*, *left-symmetric* or *chronological* algebra

Theorem 1.19 ([18]). *$PL(V)$ is the free pre-Lie algebra on the vector space V: for any pre-Lie algebra P equipped with a morphism $V \to P$, there is a unique pre-Lie morphism $PL(V) \to P$ making the following diagram commute:*

$$\begin{array}{ccc} V & \longrightarrow & PL(V) \\ & \searrow & \downarrow {\scriptstyle \exists!} \\ & & P \end{array}$$

We write PL for the free pre-Lie algebra on a space with only one element.

The free pre-Lie algebra is related to the Hopf algebra $\mathrm{H_{BCK}}$ defined in Section 1.3:

Theorem 1.20 ([18]). *The universal enveloping algebra $U(PL)$ of the free pre-Lie algebra on the one-vertex tree, viewed as a Lie algebra, is isomorphic to the dual of the Butcher–Connes–Kreimer Hopf algebra $\mathrm{H_{BCK}}$.*

In fact, the dual of the Butcher–Connes–Kreimer Hopf algebra is isomorphic to the *Grossman-Larson Hopf algebra* defined [32]. The isomorphism was proven in [38].

Pre-Lie Butcher series. Now we can formulate the pre-Lie Butcher series

Definition 1.21. A **pre-Lie Butcher series** is a formal series in $\mathbb{R}\langle \mathrm{PL} \rangle$:

$$X(\alpha) = \sum_{t \in \mathrm{PL}} h^{|t|} \alpha(t) t. \tag{1.27}$$

The classical B-series are recovered by applying the unique pre-Lie morphism associated to a vector field F:

$$\mathcal{F} : \mathrm{PL} \to \mathcal{X}(\mathbb{R}^n) \quad \text{such that} \quad \mathcal{F}(\bullet) = F.$$

This is the elementary differential function of F as defined in 1.4. It is given recursively by $\mathcal{F}(\bullet) = F$ and

$$\mathcal{F}(t) = F^{(n)}(\mathcal{F}(\tau_1), \ldots, \mathcal{F}(t_n)), \tag{1.28}$$

if $t = \mathrm{B}^+(\tau_1, \ldots, t_n)$.

B-series in any other pre-Lie algebra $(A, \triangleright)$ can be defined in the same way: by applying the unique pre-Lie algebra morphism $F : \mathrm{PL} \to A$ to the series (1.27).

Remark 1.22. Since $\mathcal{F} : \mathrm{PL} \to \mathcal{X}(\mathbb{R}^n)$ is a pre-Lie morphism, the trees associated to the derivatives of $y'(t) = F(y(t))$ can be generated by iterated grafting onto the one-vertex tree:

$$\bullet \curvearrowright (\bullet \curvearrowright (\bullet \curvearrowright \ldots (\bullet \curvearrowright \bullet) \ldots)) \quad \text{corresponds to} \quad \frac{d^n y}{dt^n}.$$

This way of looking at elementary differentials will reappear in a different setting in Chapter 2.

Remark 1.23. [Pre-Lie algebras and the Magnus expansion] The formulation of differential equations in terms of pre-Lie algebras has seen some use in numerical analysis. In [25] K. Ebrahimi-Fard and D. Manchon rephrased differential equations of the type $X'(t) = A(t)X(t)$, where X, A are linear operators in a vector space, as combinatorial equations in pre-Lie algebras. In this context they obtained an analogue of the Magnus expansion [50], a series expansion of the solution to the equation in the magma generated by monomials of pre-Lie elements. In this setting it becomes apparent that one can use the pre-Lie relation to cancel out some of the terms in the expansion, leading to a thitherto unknown reduction of the number of terms in the Magnus expansion

Chapter 2

Geometric numerical integration on manifolds

Our main objects of study in this chapter are dynamical systems evolving on *manifolds*:

$$y' = F(y), \qquad y_0 \in M, \quad F \in \mathcal{X}(M), \tag{2.1}$$

where M is a smooth manifold and $\mathcal{X}(M)$ denotes the vector fields on M. As in the previous chapter, the aim is to find good numerical approximations to the flow $\exp(tF) := \Psi_{t,F}$ of (2.1). The study of such systems comprises several different approaches: One simple way to attack the problem is to embed the manifold in $\mathbb{R}^N$, for some N, and use methods developed for $\mathbb{R}^N$ to solve the equation. But then the numerical flow of the method may drift off the manifold, and this can in some cases cause problems [28, 39, 10, 42].

A more satisfying and often better way is to use methods that are intrinsic to the manifold, and not rely on any embedding. Consider for instance a system evolving on the manifold S^3. By embedding S^3 in $\mathbb{R}^4$ one can use numerical methods that approximate the flow of the system using the basic motions of translations in $\mathbb{R}^4$. Another approach is to use *rotations* to move around S^3: $y_{n+1} = Q_n y_n$ where Q_n are orthogonal matrices, i.e. to use the action of the Lie group $SO(3)$ on S^4. This illustrates the intrinsic approach, where we are guaranteed not to drift off S^3. Methods developed for manifolds include the Crouch–Grossman and RKMK-methods (and variants thereof) [56, 57, 22, 61, 27].

In this chapter we will study a generalization of B-series called *Lie–Butcher series*. In analogy to the previous chapter we will look at the composition and substitution of Lie–Butcher series. The papers reproduced in Part II contains most of the theory and results in Lie–Butcher theory that is of interest to us here, and therefore this chapter will mainly consist of sketches of the main results, with references to the relevant papers in Part II.

2.1 Setting the stage: homogeneous manifolds and differential equations

The flows we would like to approximate evolve on smooth manifolds, and so the tools of differential geometry play an important role. We will not review the general theory of smooth manifolds here, but assume a basic knowledge of differential geometry; for excellent introductions see e.g. [67, 66]. For a viewpoint oriented toward geometric numerical integration, see [40]. More precisely, we will be working with smooth manifolds equipped with transitive actions by Lie groups, so called *homogenous manifolds*, where the Lie group provides a way to move around on the manifold.* Because the action is not in general free, the differential equation expressed on the Lie group is not in general unique. Our presentation of differential equations on homogeneous manifolds is based on the papers [59, 57, 27].

Definition 2.1. An **action of a Lie group** G on a smooth manifold M is a group homomorphism $\lambda : G \to \mathrm{Diff}(M)$, $g \mapsto \lambda_g$, where $\mathrm{Diff}(M)$ is the group of diffeomorphisms on M. We will mostly write such an action as a map $\Lambda : G \times M \to M$.

For convenience of notation we write g for the diffeomorphism λ_g, and also $g \cdot m$ for $\lambda_g(m)$. The **orbit** through a point $p \in M$ is the set $G \cdot p = \lambda_G(p)$. The action is called **transitive** if the manifold M is a single G-orbit. That is, if for all $p, q \in M$ there is a $g \in G$ so that $p = g \cdot q$. A manifold equipped with a transitive action by a Lie group G is called a **homogeneous manifold**. A consequence of this is that M is diffeomorphic to the right cosets G/G_x of G, where G_x is the closed Lie subgroup of isotropies, $G_x = \{g \in G \mid gx = x\}$ (the point stabilizer): the smooth manifold structure of G/G_x comes from the quotient map, and the diffeomorphism $F : G/G_x \to M$ is given by $F(gG_x) = g \cdot x$. The group G_x is called *the* subgroup of isotropies because if x' is another point in G, then G_x and $G_{x'}$ are conjugate, and therefore isomorphic.

Important examples of homogeneous manifolds are the spheres $S^n = SO(n+1)/SO(n)$. A (somewhat degenerate) example is the homogeneous manifold $(\mathbb{R}^n, (\mathbb{R}^n, +))$. Here the action of $\mathbb{R}^n$ on itself is given by translations. The theory developed for homogeneous manifolds in this chapter will reduce to the theory developed in the previous chapter when applied to this particular case.

Actions by Lie groups on manifolds can be associated to actions by Lie algebras. Let $\Lambda : G \times M \to M$ be an action of G on M. The associated Lie algebra action $\lambda_* : \mathfrak{g} \to \mathcal{X}(M)$ of $\mathfrak{g}$ on M is the homomorphism defined by:

$$\lambda_*(v)(p) = \frac{d}{dt}\bigg|_{t=0} \Lambda(\exp(tV), p). \tag{2.2}$$

* Note that other manifolds with *local* actions could also be considered, but to to avoid unnecessary complications we elect to only consider homogeneous manifolds.

We sometimes write $v \cdot y$ for the element $\lambda_*(v)(y) \in T_yM$. The *Lie–Palais theorem* [62] ensures us that as long as the Lie group G is simply connected, then every action by $\mathfrak{g}$ comes from an action by G. However, if the Lie group is not simply connected, then we can only lift the $\mathfrak{g}$-action to the universal covering group of G. If $F \in \mathcal{X}(M)$ is a vector field, then an element v so that $\lambda_*(v) = F$ is called an *infinitesimal generator* for F.

Remark 2.2. In some cases it makes sense to use other maps $\phi : \mathfrak{g} \to G$ (satisfying $\phi(0) = e$ and $\phi'(0) = V$) besides the exponential map to construct maps $\mathfrak{g} \to \mathcal{X}(M)$ as in Equation (2.2). An overview of various maps of this kind, and their usefulness, can be found in [27].

Differential equations on homogeneous manifolds. Consider the differential equation on a homogeneous manifold (M, G, λ):

$$y' = F(y), \qquad y_0 \in M, \quad F : M \to TM. \tag{2.3}$$

The solution is the flow $\Psi_{t,F} = \exp(tF)$ of the vector field F. The vector field can be written in terms of its infinitesimal generator as $F = \lambda_*(v) : M \to TM$ for an element $v \in \mathfrak{g}$, and the transitivity of the action also allows us to construct a map $f : M \to \mathfrak{g}$ so that

$$F(y) = \lambda_*(f(y))(y) = f(y) \cdot y \tag{2.4}$$

Note that as long as the action is not free, this f is not unique: if $f : M \to \mathfrak{g}$ is such a map, then $f + i : M \to \mathfrak{g}$, where $i(p)$ is in the isotropy subalgebra $\mathfrak{g}_p$ of $\mathfrak{g}$, is another map of the same type. This choice of isotropy class can be helpful when constructing numerical integrators [46].

The differential equation (2.3) can be written as:

$$y' = f(y) \cdot y, \quad \text{where} \quad f : M \to \mathfrak{g}, \tag{2.5}$$

and this is the type of differential equation we will consider in this chapter. Note that in the classical case of $(\mathbb{R}^n, (\mathbb{R}^n, +))$, this equation reduces to the ordinary differential equation (2.3). We also note that the class contains the equations formulated in terms of *frames*:

Remark 2.3 (Frames and differential equations)**.** In the literature for numerical integration of differential equations on manifolds the equations are often simplified by using a *frame* on the manifold [61, 60, 15]. A frame is a set of vector fields $\{E_i\}$ that at each point on the manifold spans the tangent space at that point, so that any vector field F can be written as $F = \sum_i f_i E_i$. The flow equation (2.3) for F can then be written as

$$y' = \sum_i f_i(y) E_i(y), \quad \text{where} \quad f_i : M \to \mathbb{R} \quad \text{are smooth.} \tag{2.6}$$

If we write $\mathfrak{g} \subset \mathcal{X}(M)$ for the Lie subalgebra generated by the vector fields $\{E_i\}$, and let $\lambda_* : \mathfrak{g} \to \mathrm{Diff}(M)$ be as in (2.2), we see that Equation (2.6) is a special case of Equation (2.5), with $f : M \to \mathfrak{g}$ defined by $f(y) = \sum_i f_i(y) E_i$.

Remark 2.4. In [27], K. Engø formulated the general operation of 'moving' differential equations between manifolds using equivariance of actions and relatedness of vector fields. In particular, every differential equation of the form (2.5) was shown to be equivalent to a differential equation on $\mathfrak{g}$. The following diagram from [27] summarizes this:

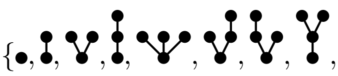
$$\begin{array}{ccccc} T\mathfrak{g} & \xrightarrow{T(\exp)} & TG & \xrightarrow{T(\lambda.(p))} & TM \\ \uparrow & & \uparrow & & \uparrow \lambda_*(v)(p) \\ \mathfrak{g} & \xrightarrow[\exp]{} & G & \xrightarrow[\lambda.(p)]{} & M \end{array}$$

In other words, the differential equation on a homogeneous manifold (M, G) is moved to the Lie group G (the middle vertical arrow) and then to the Lie algebra $\mathfrak{g}$ (the first vertical arrow). As before, the exponential map $\exp : \mathfrak{g} \to G$ can in some cases be replaced by other maps. The construction of the vertical arrows can be found in [27]. This is the result exploited in the so-called RKMK methods [55, 56, 57].

2.2 Trees, D-algebras and Lie–Butcher series

In Chapter 1 we observed that ordinary differential equations in $\mathbb{R}^n$ are related to rooted trees, and that the formal series indexed over trees we used in our study are related to pre-Lie algebras. In the more general case of differential equations on manifolds, we will see that forests of *ordered* rooted trees and D-algebras play these roles. We will sketch the construction of ordered rooted trees, D-algebras and Lie–Butcher series. Details can be found in [58] or [47] (Paper A in Part II below).

Ordered trees and D-algebras. The set

OT = {  , . . . }.

of ordered rooted trees consists of *all* rooted trees (Section 1.2). Unlike the set $\mathrm{T} \subset \mathrm{OT}$ of rooted trees, we do not identify trees who differ in the order of their branches. In other words, an ordered rooted tree is a tree τ together with a chosen order of the branches connected to each vertex of τ. Write OF for the set of ordered words (including the empty word) of elements from OT, called the set of **ordered forests**. Let $\mathrm{N} = \mathbb{R}\langle\mathrm{OT}\rangle$ be the noncommutative polynomials over OT. The linear dual $\mathrm{N}^* := \mathrm{Hom}(\mathrm{N}, \mathbb{R})$ is identified with the infinite combinations of words, and we write $\langle\cdot,\cdot\rangle$ for the pairing making words in OT orthogonal. That is, $\langle\omega_1, \omega_2\rangle = \delta_{\omega_1,\omega_2}$, for all $\omega_1, \omega_2 \in \mathrm{OF}$.

It is sometimes convenient to allow the trees to be *decorated* by a set $\mathcal{C}$, often called the set of colors. This is done via a map from the vertices of the tree to the set $\mathcal{C}$. We write $\mathrm{OT}_{\mathcal{C}}$ and $\mathrm{OF}_{\mathcal{C}}$ for the set of trees and forests colored by $\mathcal{C}$.

A basic operation on N is the **left grafting product** $\cdot \curvearrowright \cdot : \mathrm{N} \otimes \mathrm{N} \to \mathrm{N}$ of [58]. It is defined recursively by

$$\begin{aligned}
&\mathbb{I} \curvearrowright \omega = \omega \\
&\omega \curvearrowright \mathbb{I} = 0 \\
&\omega \curvearrowright \bullet = B^{+}(\omega), \\
&\tau \curvearrowright \omega_1\omega_2 = (\tau \curvearrowright \omega_1)\omega_2 + \omega_1(\tau \curvearrowright \omega_2) \\
&(\tau\omega) \curvearrowright \omega_1 = \tau \curvearrowright (\omega \curvearrowright \omega_1) - (\tau \curvearrowright \omega) \curvearrowright \omega_1,
\end{aligned} \tag{2.7}$$

where τ is a tree and ω_1, ω_2 are forests. If we write $(\cdot)[\cdot]$ for $\curvearrowright$, then concatenation and grafting gives N the structure of a D-algebra, as defined in [58] (see also [47, 49, 48]):

Definition 2.5. Let A be a unital associative algebra with product $f, g \mapsto fg$, unit $\mathbb{I}$ and equipped with a non-associative composition $(.)[.] : A \otimes A \to A$ such that $\mathbb{I}[g] = g$ for all $g \in A$. Write $\mathcal{D}(A)$ for the set of all $f \in A$ such that $f[\cdot]$ is a derivation:

$$\mathcal{D}(A) = \{f \in A \mid f[gh] = (f[g])h + g(f[h]) \text{ for all } g, h \in A\}.$$

Then A is called a **D-algebra** if for any derivation $f \in \mathcal{D}(A)$ and any $g \in A$ we have

(i) $g[f] \in \mathcal{D}(A)$
(ii) $f[g[h]] = (fg)[h] + (f[g])[h]$.

In [58] it was also shown that the D-algebra N is the *free* D-algebra:

Theorem 2.6 ([58])**.** *The vector space* $\mathrm{N} = \mathrm{k}\langle \mathrm{OT}_{\mathcal{C}} \rangle$ *is the free D-algebra over* $\mathcal{C}$*. That is, for any D-algebra* $\mathcal{A}$ *and any map* $\nu \colon \mathcal{C} \to D(\mathcal{A})$ *there exists a unique D-algebra homomorphism* $\mathcal{F}_\nu \colon N \to \mathcal{A}$ *such that* $\mathcal{F}_\nu(c) = \nu(c)$ *for all* $c \in \mathcal{C}$*.*

$$\begin{array}{ccc}
\mathcal{C} & \hookrightarrow & \mathrm{N} \\
\nu \downarrow & & \downarrow \exists!\, \mathcal{F}_\nu \\
D(\mathcal{A}) & \hookrightarrow & \mathcal{A}
\end{array}$$

A D-algebra homomorphism between two D-algebras A and B is an algebra morphism $F : A \to B$ such that $F(\mathcal{D}(A)) \subset \mathcal{D}(B)$, and $F(a[b]) = F(a)[F(b)]$.

This theorem enables us to define elementary differentials and Lie–Butcher series by applying it to the case where $\mathcal{A}$ is the D-algebra $U(\mathfrak{g})$ of differential operators. Recall that a vector field (or, in other words, a first-order differential operator) F on a homogeneous manifold (M, G) can be represented as a function $f : M \to \mathfrak{g}$. Similarly, all higher order differential operators on M can be represented as functions from M to the universal enveloping algebra $U(\mathfrak{g})$ of $\mathfrak{g}$.

Theorem 2.7 ([58]). *Let (M, G) be a homogeneous manifold and let $\mathfrak{g}$ denote the Lie algebra of G. Let $U(\mathfrak{g})$ denote the universal enveloping algebra of $\mathfrak{g}$, consisting of all higher order differential operators on M, and extend its structure to $C^\infty(M, U(\mathfrak{g})) =: U(\mathfrak{g})^M$ via*

$$F[G](p) := (F(p)[G])(p), \quad FG(p) := F(p)G(p). \tag{2.8}$$

These two operations give $U(\mathfrak{g})^M$ the structure of a D-algebra.

Remark: post-Lie algebras. In [48] (reproduced as Paper C in Part II) the author and H. Munthe-Kaas developed a more refined view of D-algebras, where the D-algebras are enveloping algebras of *post-Lie algebras* (post-Lie algebras were also introduced independently by Vallette in [70]). This point of view is currently being studied further in an ongoing project [24], where the *operad* behind post-Lie and D-algebras (also called **post associative algebras**) is explored.

Definition 2.8. A **post-Lie algebra** is a Lie algebra $(A, [\cdot,\cdot])$ equipped with a non-commutative, non-associative product $\rhd : A \otimes A \to A$ satisfying:

$$x \rhd [y,z] = [x \rhd y, z] + [y, x \rhd z] \qquad \text{(derivation property)} \tag{2.9}$$

$$[x,y] \rhd z = a_\rhd(x,y,z) - a_\rhd(y,x,z), \tag{2.10}$$

where $a_\rhd(x,y,z)$ is the associator $a_\rhd(x,y,z) = x \rhd (y \rhd z) - (x \rhd y) \rhd z$.

In [48] it is shown that the free Lie algebra over rooted trees colored by a set $\mathcal{C}$ is the free post-Lie algebra, and that its universal enveloping algebra is the free D-algebra defined above. Notice that relation (2.10) implies that a pre-Lie algebra (Section 1.5) is a post-Lie algebra with vanishing bracket.

Lie–Butcher series

Analogous to the B-series of Chapter 1, the Lie–Butcher series can be used to represent flows – numerical or exact – on homogeneous manifolds. To achieve this one combines the concept of *Lie series* in free Lie algebras with ideas from the theory of B-series. An exposition of free Lie algebras and Lie series can be found in the book [63] by Reutenauer.

The **free Lie algebra** $\mathrm{FLA}(A)$ over a set A of generators is the closure of the generators under commutation and linear combination. In particular, we have the free Lie algebra $\mathrm{FLA}(\mathrm{OT})$ over the set of ordered rooted trees. A **Lie series** is a series expansion:

$$S = \sum_{n \geq 0} S_n, \tag{2.11}$$

where each homogeneous component is an element of $\mathrm{FLA}(\mathrm{OT})$, i.e. the S_n's are *Lie polynomials*.

A Lie series of particular interest to us appears when computing the pullback of functions along flows of vector fields on homogeneous manifolds. Let $F \in \mathcal{X}(M)$ be a vector field with flow $\Phi_{t,F}$, and $\psi : M \to \mathfrak{g}$ a function. Then

$$\left.\frac{d}{dt}\right|_{t=0} \Phi^*_{t,F}\psi = F[\psi]. \tag{2.12}$$

The Taylor expansion of $\Phi^*_{t,F}\psi$ around 0 therefore takes the form of a Lie series

$$\begin{aligned} \Phi^*_{t,F}\psi &= \sum_{n=0}^{\infty} \frac{t^n}{n!} \left(\left.\frac{\partial^n}{\partial t^n}\right|_{t=0} \Phi^*_{t,F}\psi \right) \\ &= \psi + tF[\psi] + \frac{t^2}{2!}F[F[\psi]] + \frac{t^3}{3!}F[F[F[\psi]]] + \cdots . \end{aligned} \tag{2.13}$$

Bell polynomials. The higher order derivatives of the pullbacks can be written in terms of noncommutative Bell polynomials [47]:

Definition 2.9. Let $D = \mathbb{R}\langle \mathcal{I} \rangle$ be the free associative algebra over an alphabet $\mathcal{I} = \{d_i\}$, and let $\partial : D \to D$ denote the derivation given by $\partial(d_i) = d_{i+1}$. The **noncommutative Bell polynomials** $B_n = B_n(d_1, \dots, d_n) \in \mathbb{R}\langle \mathcal{I} \rangle$ are defined by the recursion

$$\begin{aligned} B_0 &= \mathbb{I} \\ B_n &= (d_1 + \partial)B_{n-1}, \quad n > 0. \end{aligned} \tag{2.14}$$

Theorem 2.10 ([55, 47])**.** *The derivatives of the pullback of a function ψ along the time-dependent flow $\Phi_{t,F}$ is:*

$$\frac{d^n}{dt^n}\Phi^*_{t,F}\psi = B_n(F)[\psi], \tag{2.15}$$

where $B_n(F_t)$ is the image of the Bell polynomials B_n under the homomorphism given by $d_i \mapsto F^{(i-1)}$ ($(i-1)$th derivative). In particular

$$\left.\frac{d^n}{dt^n}\right|_{t=0} \Phi^*_{t,F_t}\psi = B_n(F_1, \dots, F_n)[\psi] =: B_n(F_i)[\psi], \tag{2.16}$$

where $F_{n+1} = d^n/dt^n|_{t=0}F$.

This result allows us to rewrite the Lie series (2.13) as the following expression [55]:

$$\Phi^*_{t,F}\psi = \sum_{n=0}^{\infty} F^n[\psi]\frac{t^n}{n!} = \sum_{n=0}^{\infty} B_n(F_i)[\psi]\frac{t^n}{n!}, \tag{2.17}$$

where F^n iterated application of F, as in Equation (2.13).

Remark 2.11. It is well known that the classical Bell polynomials can be defined in terms of determinants, and it seems like the non-commutative Bell polynomials can be defined in the same way, only now in terms of a non-commutative analog of the determinant: the **quasi-determinants** of Gelfand and Retakh ([30], see also [29]). For example, we have

$$\begin{aligned}
\det \begin{bmatrix} x_1 & -1 & 0 \\ \binom{3-1}{1}x_2 & x_1 & -1 \\ \binom{3-1}{2}x_3 & \binom{3-2}{1}x_2 & x_1 \end{bmatrix} &= \det \begin{bmatrix} x_1 & -1 & 0 \\ 2x_2 & x_1 & -1 \\ x_3 & x_2 & x_1 \end{bmatrix} \\
&= x_1^3 + 2x_1x_2 + x_2x_1 + x_3 \\
&= B_3,
\end{aligned}$$

where det denotes the quasi-determinant. The significance of this result is at the present time unexplored.

The Lie–series (2.13) can also be written as the *Lie–Butcher series* for the exact flow.

Lie–Butcher series. The general Lie–Butcher series $\mathcal{B}_f(\alpha)$ are constructed to represent flows given by $y_0 \mapsto y_t = \Psi_t(y_0)$:

$$\Psi_t(y(t)) = \mathcal{B}_f(\alpha)[\Psi_t](y_0). \tag{2.18}$$

Before giving the definition of Lie–Butcher series we need to define the elementary differentials of a vector field F:

Definition 2.12. Let $\mathcal{F}_f : \mathrm{N} \to U(\mathfrak{g})^M$ be the unique D-algebra morphism given by Theorem 2.6 by associating $\bullet$ to a vector field $f : M \to \mathfrak{g}$. This is called the **elementary differentials** of the vector field f.

Note that $\mathcal{F}_f : \mathrm{N} \to U(\mathfrak{g})^M$ is given recursively by

(i) $\mathcal{F}_f(\mathbb{I}) = \mathbb{I}$

(ii) $\mathcal{F}_f(\mathrm{B}^+(\omega)) = \mathcal{F}_f(\omega)[f]$

(iii) $\mathcal{F}_f(\omega_1\omega_2) = \mathcal{F}_f(\omega_1)\mathcal{F}_f(\omega_2)$

The general Lie–Butcher series are expansions of elementary differentials indexed over ordered rooted forests.

Definition 2.13. A **Lie–Butcher series** (LB-series) is a formal series expansion in $U(\mathfrak{g})^M$:

$$\mathcal{B}_f(\alpha) = \sum_{\omega \in \mathrm{OF}} h^{|\omega|} \alpha(\omega) \mathcal{F}_f(\omega), \tag{2.19}$$

where $\alpha : \mathrm{N} \to \mathbb{R}$.

It turns out [47] that the Lie series (2.13) can be written as

$$\Phi^*_{t,f} \psi = \sum_{\omega \in \mathrm{OT}} \gamma(\omega) \mathcal{F}_f(\omega), \tag{2.20}$$

where γ are the coefficients appearing when iteratively (left) grafting $\bullet$ onto $\bullet$. This is the Lie–Butcher series for the exact flow.

See [55, 56, 61, 60, 58], Paper A [47] and Paper B [49] in Part II for examples of and details about LB-series and numerical flows.

2.3 Composition of Lie–Butcher series

We would like to understand the result of *composing* LB-series methods in a similar way as we did for B-series methods in Section 1.3. The basic problem is to determine whether the method Φ resulting from composing two methods $\Phi^2 \circ \Phi^1$– both given by LB-series–is another LB-series, and in that case, what its coefficients are. Just as there is a Hopf algebra governing composition of B-series (the BCK Hopf algebra discussed in Section 1.3), there is a Hopf algebra $\mathrm{H}_{\mathrm{MKW}}$ behind the composition of LB-series. This Hopf algebra was first studied in [58], where its properties and its relation to the BCK Hopf algebra was explored. An introduction can also be found in [47], reproduced as Paper A in Part II.

The Hopf algebra of composition. As a vector space $\mathrm{H}_{\mathrm{MKW}}$ is spanned by the set of ordered forests: $\mathrm{H}_{\mathrm{MKW}} = \mathbb{R}\langle\mathrm{OT}\rangle$. The product is given by *shuffling*:

$$\begin{aligned} &\mathbb{I} \sqcup\!\sqcup\, \omega = \omega = \omega \sqcup\!\sqcup\, \mathbb{I} \\ &(\tau_1\omega_1) \sqcup\!\sqcup\, (\tau_2\omega_2) = \tau_1(\omega_1 \sqcup\!\sqcup\, \tau_2\omega_2) + \tau_2(\tau_1\omega_1 \sqcup\!\sqcup\, \omega_2) \end{aligned} \tag{2.21}$$

where $\tau_1, \tau_2 \in \mathrm{OT}$ and $\omega_1, \omega_2 \in \mathrm{OF}$. The coproduct is given recursively by $\Delta_N(\mathbb{I}) = \mathbb{I} \otimes \mathbb{I}$ and

$$\Delta_N(\omega\tau) = \omega\tau \otimes \mathbb{I} + \Delta_N(\omega) \sqcup\!\sqcup \cdot (I \otimes B_i^+) \Delta_N(B^-(\tau)), \tag{2.22}$$

where $\tau \in \mathrm{OT}$, $\omega \in \mathrm{OF}$. Here $\sqcup\!\sqcup \cdot : \mathrm{N}^{\otimes 4} \to \mathrm{N} \otimes \mathrm{N}$ denotes shuffle on the left and concatenation on the right: $(\omega_1 \otimes \omega_2) \sqcup\!\sqcup \cdot (\omega_3 \otimes \omega_4) = (\omega_1 \sqcup\!\sqcup\, \omega_3) \otimes (\omega_2\omega_4)$.

The coproduct can also be written in terms of *left admissible cuts*, analogous to the coproduct in $\mathrm{H}_{\mathrm{BCK}}$ (Theorem 1.12):

Theorem 2.14 ([58]). *The coproduct in* $\mathrm{H}_{\mathrm{MKW}}$ *can be written as*

$$\Delta_{MKW}(\omega) = \sum_{c \in FLAC(\omega)} P^c(\omega) \otimes R^c(\omega), \tag{2.23}$$

where ω *is a forest in* OT.

A left admissible cut differs from the admissible cuts defined in Section 1.3 (see [58]): an *elementary cut* c of a tree τ is a selection of edges to be removed from τ, chosen in such a way that if an edge e is removed, then all the branches on the same level and to the left of e must also be removed. A cut results in a collection of trees concatenated together to form a forest $P^c_{el}(\tau)$ (the *pruned part*), and a remaining tree $R^c_{el}(\tau)$, containing the root. A left admissible cut $c = \{c_1, \ldots, c_n\}$ on τ is a collection of such elementary cuts, with the property that any path from the root to any vertex crosses at most one cut c_i. The pruned parts from each cut together form the pruned part $P^c(\tau)$ of the left admissible cut, where the parts coming from different cuts are shuffled together. We also include the *full cut* and the *empty cut*, which results in $P^c(\tau) = \tau$ and $P^c(\tau) = \mathbb{I}$, respectively. The cutting operation is extended to forests ω as follows: apply the B^+ operation to ω to get a tree, cut this without using cuts of edges coming out of the root, and, finally, remove the added root from $R^c(\omega)$.

See Table 2.1 for some examples of the coproduct Δ_{MKW}, and see [58] or [47] (reproduced as Paper A in Part II below) for further examples and other properties of $\mathrm{H}_{\mathrm{MKW}}$.

ω	$\Delta_{MKW}(\omega)$
$\mathbb{I}$	$\mathbb{I} \otimes \mathbb{I}$
$\bullet$	$\bullet \otimes \mathbb{I} + \mathbb{I} \otimes \bullet$
$\bullet\bullet$	$\bullet\bullet \otimes \mathbb{I} + \bullet \otimes \bullet + \mathbb{I} \otimes \bullet\bullet$
$[\bullet]$	$[\bullet] \otimes \mathbb{I} + \bullet \otimes \bullet + \mathbb{I} \otimes [\bullet]$
$\bullet[\bullet]$	$\bullet[\bullet] \otimes \mathbb{I} + 2\,\bullet\bullet \otimes \bullet + \bullet \otimes [\bullet] + \bullet \otimes \bullet\bullet + \mathbb{I} \otimes \bullet[\bullet]$
$[\bullet]\bullet$	$[\bullet]\bullet \otimes \mathbb{I} + [\bullet] \otimes \bullet + \bullet \otimes \bullet\bullet + \mathbb{I} \otimes [\bullet]\bullet$

Table 2.1: Examples of the coproduct Δ_{MKW}

The main result linking $\mathrm{H}_{\mathrm{MKW}}$ to LB-series is the following, which is an analog of the Hairer-Wanner theorem (Theorem 1.13) for B-series:

Theorem 2.15 ([58]). *The composition of two LB-series is again a LB-series:*

$$\mathcal{B}_f(\alpha)[\mathcal{B}_f(\beta)] = \mathcal{B}_f(\alpha * \beta), \tag{2.24}$$

where $*$ *is the convolution product in* $\mathrm{H}_{\mathrm{MKW}}$.

2.4 Substitution and backward error analysis for Lie–Butcher series

In [49] (reproduced as Paper B in Part II) the substitution law for LB-series methods was developed, culminating in a formula that can be used to calculate the modified vector field used in backward error analysis.

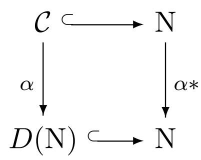
The substitution law. The basic idea is as for B-series (Section 1.4): We consider substituting a LB-series into another LB-series, e.g. $\mathcal{B}_{\mathcal{B}_f(\beta)}(\alpha)$, and the question is as before: is this a LB-series, and in that case, which one? The result is given in terms of the *substitution law*:

Theorem 2.16 ([49])**.** *The substitution law defined in Definition 2.17 corresponds to the substitution of LB-series in the sense that*

$$\mathcal{B}_{\mathcal{B}_f(\beta)}(\alpha) = \mathcal{B}_f(\beta \star \alpha)$$

The substitution law is defined by using the freeness of the D-algebra $\mathrm{N} = \mathbb{R}\langle\mathrm{OT}\rangle$ (Theorem 2.6):

Definition 2.17. For any map $\alpha : \mathcal{C} \to D(\mathrm{N})$ Theorem 2.6 implies that there a unique D-algebra homomorphism $\alpha* : \mathrm{N} \to \mathrm{N}$ such that $\alpha(c) = \alpha * c$ for all $c \in \mathcal{C}$. This homomorphism is called α-substitution.



Calculating the substitution law. To obtain a formula for the substitution law, we consider the dual α_*^t of α-substitution:

$$\langle \alpha * \beta, \omega \rangle = \langle \beta, \alpha_*^t(\omega) \rangle, \tag{2.25}$$

and we call it the *substitution character*. The dual pairing $\langle \cdot, \cdot \rangle$ is the one induced by requiring that all forests in OT are orthogonal, and we may write $\langle \alpha, \omega \rangle = \alpha(\omega)$. The map α_*^t is a character for the shuffle product [49, Proposition 3.8]: $\alpha_*^t(\omega_1 \sqcup\!\sqcup\, \omega_2) = \alpha_*^t(\omega_1) \sqcup\!\sqcup\, \alpha_*^t(\omega_2)$.

The formula for the substitution law is based on the cutting of trees as in the coproduct Δ_{MKW}. More specifically, it is based on the dual of grafting, called *pruning*:

$$\mathcal{P}_\nu(\omega) = \sum_{c \in LAC(\omega)} \langle \nu, P^c(\omega) \rangle R^c(\omega). \tag{2.26}$$

Here the sum is over the left admissible cuts, but as opposed to the cuts in the formula (2.23) for Δ_{MKW}, the full cut is not included.

In [49] the following inductive formula for α_*^t was obtained:

Theorem 2.18 ([49]). *We have*

$$\alpha_*^t(\omega) = \sum_{(\omega)\in\Delta_C} \sum_{c\in LAC(\omega_{(2)})} \alpha_*^t(\omega_{(1)})\, \mathrm{B}^+\left(\alpha_*^t(P^c(\omega_{(2)}))\right) \alpha(R^c(\omega_{(2)})),$$

if $\omega \neq 1$ *and* $\alpha_*^t(\mathbb{I}) = \mathbb{I}$. *Here* Δ_C *denotes the deconcatenation coproduct.*

By introducing a magmatic operation $\mu_\times$ on N, given by $\mu_\times(\omega_1, \omega_2) = \omega_1\, \mathrm{B}^+(\omega_2)^\dagger$, this can also be written as a composition of operators:

$$\alpha_*^t = \mu \circ (\mu_\times \otimes I) \circ (\alpha_*^t \otimes \alpha_*^t \otimes a) \circ (I \otimes \Delta'_{MKW}) \circ \Delta_C. \tag{2.27}$$

Here Δ_C is deconcatenation, Δ'_{MKW} denotes the coproduct in (2.23) with the full cut removed, and μ denotes concatenation.

Some examples of the substitution character can be found in Table 2.2. Many more examples and details can be found in [49] (Paper B below).

ω	$\alpha_*^t(\omega)$
$\mathbb{I}$	$\mathbb{I}$
$\bullet$	$\alpha(\bullet)\bullet$
$\bullet\bullet$	$\alpha(\bullet)^2\bullet\bullet$
$B^+(\bullet)$	$\alpha(B^+(\bullet))\bullet + \alpha(\bullet)^2 B^+(\bullet)$
$\bullet B^+(\bullet)$	$\alpha(\bullet B^+(\bullet))\bullet + \alpha(\bullet)\alpha(B^+(\bullet))\bullet\bullet + \alpha(\bullet)^3 \bullet B^+(\bullet)$
$B^+(\bullet)\bullet$	$\alpha(B^+(\bullet)\bullet)\bullet + \alpha(\bullet)\alpha(B^+(\bullet))\bullet\bullet + \alpha(\bullet)^3 B^+(\bullet)\bullet$

Table 2.2: Examples of the substitution character α_*^t

Remark 2.19. One would like the substitution law $*$ to be a convolution product in a Hopf or bialgebra, analogous to the substitution of B-series (Theorem 1.14). One possible way to achieve this is by obtaining a concrete description of the operations in the post-Lie operad. In that case one can follow the procedure in [9], which, roughly, is the following: The post-Lie operad has a pre-Lie structure (general phenomenon for augmented operads), there is an associated Lie algebra structure, its universal enveloping algebra is a Hopf algebra, and its dual is the Hopf algebra for the substitution law. This is a project currently under investigation [24].

† This magmatic operation $\mu_\times$ allows us to rewrite all the basic operations of Lie–Butcher theory in a simpler way, a way which is also convenient for implementation. See Paper B ([49]) for details.

Chapter 3

Summaries of papers

Summary of Paper A

Hopf algebras of formal diffeomorphisms and numerical integration on manifolds

A. Lundervold and H.Z. Munthe-Kaas

Published in *Contemporary Mathematics,* volume 539, 2011

This paper explores several of the algebraic structures appearing in the study of Lie group integrators: Hopf algebras, Lie series, Lie-Butcher series, Lie idempotents, a noncommutative Faà di Bruno algebra and noncommutative Bell polynomials. It serves both as an introduction to relevant algebraic concepts for numerical analysts, and as an introduction to numerical analysis for algebraists. It is partly a review and partly a research paper. Some of the results in the paper can be found elsewhere in the literature; others are original.

Among other things, the paper gives a purely algebraic way to understand Lie–Butcher theory, in the spirit of the paper [58] by H. Munthe-Kaas and W. Wright. The theory is formulated in terms of the ordered rooted trees OT, together with a few basic operations making it a D-algebra (Section 2.2, Part I). Various representation of flows written in terms of Lie–Butcher series are discussed, and we find algebraic methods for converting between the representations. This involves Lie idempotents and the non-commutative Bell polynomials (slightly reformulated to give an operator we call Q):

Flows $y_0 \mapsto y(t) = \Psi_t(y_0)$ on a homogeneous manifold M can be represented by LB-series in several different ways:

1. In terms of pullback series: Find a character α in $\mathrm{H}_{\mathrm{MKW}}$ such that

$$\Psi(y(t)) = \mathcal{B}_t(\alpha)(y_0)[\Psi] \quad \text{for any } \Psi \in U(\mathfrak{g})^M. \tag{3.1}$$

2. In terms of an autonomous differential equation: Find an infinitesimal character β in $\mathrm{H}_{\mathrm{MKW}}$ such that $y(t)$ solves

$$y'(t) = \mathcal{B}_t(\beta)(y(t)). \tag{3.2}$$

3. In terms of a non-autonomous equation of *Lie type*: Find an infinitesimal character γ in H_{sh} such that $y(t)$ solves

$$y'(t) = \left(\frac{\partial}{\partial t}\mathcal{B}_t(\gamma)(y_0)\right) y(t). \tag{3.3}$$

The relationships between the coefficients α, β and γ in the above LB-series can be expressed as follows:

$\beta = \alpha \circ e$	e is the eulerian idempotent in $\mathrm{H}_{\mathrm{MKW}}$.
$\alpha = \exp^{\diamond}(\beta)$	Exponential wrt. GL-product
$\gamma = \alpha \circ Y^{-1} \circ D$	Dynkin idempotent in $\mathrm{H}_{\mathrm{sh}}(\mathrm{OT})$.
$\alpha = Q(\gamma)$	Q-operator in $\mathrm{H}_{\mathrm{sh}}(\mathrm{OT})$.

Here H_{sh} denotes the shuffle Hopf algebra, and Q is constructed from the Bell polynomials.

Summary of Paper B

Backward error analysis and the substitution law for Lie group integrators

A. Lundervold and H.Z. Munthe-Kaas

Submitted to *Foundations of Computational Mathematics,* 2011.

Paper A ends with a short presentation of the substitution law for Lie–Butcher series, which Paper B develops in full detail. We obtain a formula for the substitution law that can be used to calculate the coefficients of the modified vector fields used in backward error analysis.

The paper continues in the tradition of Paper A by explaining how Lie–Butcher theory is purely algebraic. For example, it points out how all the basic definitions follow from the fact that $\mathrm{N} = \mathbb{R}\langle OT \rangle$ (as defined in Section 2.2 in Part I) is the *free* D-algebra. Then elementary differentials F_f, Lie–Butcher series $\mathcal{B}_f$ and also the substitution law $\star$ can be defined in terms of commutative diagrams:

$$\begin{array}{ccc} \{\bullet\} & \hookrightarrow & \mathrm{N} \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \mathcal{F}_f} \\ \mathfrak{g}^M & \hookrightarrow & U(\mathfrak{g})^M \end{array} \qquad \begin{array}{ccc} \{\bullet\} & \hookrightarrow & \mathrm{N}^* \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \mathcal{B}_f} \\ \mathfrak{g}^M & \hookrightarrow & U(\mathfrak{g})^M \end{array} \qquad \begin{array}{ccc} \{\bullet\} & \hookrightarrow & \mathrm{N} \\ {\scriptstyle a}\downarrow & & \downarrow{\scriptstyle a\star} \\ D(\mathrm{N}) & \hookrightarrow & \mathrm{N}\,. \end{array}$$

A future goal will be to describe the Hopf algebra underlying the substitution law, a project currently under investigation [24].

Summary of Paper C

On pre-Lie-type algebras with torsion

A. Lundervold and H.Z. Munthe-Kaas

The main motivation for this paper comes from the observation that pre-Lie algebras correspond to algebras of affine connections with vanishing curvature and torsion, which is reflected in their use in classical geometric numerical integration in $\mathbb{R}^n$. As we have seen, the role of pre-Lie algebras are taken over by D-algebras when we look at geometric numerical integration on more general manifolds, which may include both curvature and torsion. In this paper we introduce an algebraic formulation for the case of connections with non-vanishing curvature or torsion.

	flat	const. curvature
torsion-free	**PreLie**	**Lie admissible**
const. torsion	**PostLie**	?*

It turns out that the correct algebraic formulation for flat algebras with constant torsion is **post Lie algebras**. This paper relates these to the D-algebras of numerical integration by showing how the universal enveloping algebra of the free post-Lie algebra is isomorphic to the free D-algebra. This opens up a new way to study Lie–Butcher series, more closely related to their character as "Lie-series". It also gives a cleaner way to understand their geometric features.

* The case corresponding to constant curvature and torsion has not yet been discovered.

Bibliography

[1] E. Abe. *Hopf Algebras*. Cambridge University Press, 1980.

[2] A.A. Agrachev and R.V. Gamkrelidze. Chronological algebras and nonstationary vector fields. *Journal of Mathematical Sciences*, 17(1):1650–1675, 1981.

[3] V.I. Arnold. *Mathematical methods of classical mechanics*. Springer, second edition, 1989.

[4] C. Brouder. Runge-Kutta methods and renormalization. *The European Physical Journal C: Particles and Fields*, 12(3):521–534, 2000.

[5] C.J. Budd and M.D. Piggott. Geometric integration and its applications. In P.G. Ciarlet and F. Cucker, editors, *Handbook of numerical analysis: Foundations of Computational Mathematics*, volume 11, pages 35–139. Elsevier, 2003.

[6] J.C. Butcher. Coefficients for the study of Runge-Kutta integration processes. *Journal of the Australian Mathematical Society*, 3(02):185–201, 1963.

[7] J.C. Butcher. An algebraic theory of integration methods. *Mathematics of Computation*, 26(117):79–106, 1972.

[8] J.C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons Inc, second edition, 2008.

[9] D. Calaque, K. Ebrahimi-Fard, and D. Manchon. Two interacting Hopf algebras of trees: A Hopf-algebraic approach to composition and substitution of B-series. *Advances in Applied Mathematics*, 47(2), 2011.

[10] M. Calvo, A. Iserles, and A. Zanna. Runge–Kutta methods on manifolds. In D.F. Griffiths and G.A. Watson, editors, *Numerical Analysis, A.R. Mitchell 75th Birthday Volume*, pages 57–70. World Scientific, 1996.

[11] M.P. Calvo, A. Murua, and J.M. Sanz-Serna. Modified equations for ODEs. *Contemporary Mathematics*, 173:63–74, 1994.

[12] M.P. Calvo and J.M. Sanz-Serna. Canonical B-series. *Numerische Mathematik*, 67(2):161–175, 1994.

[13] P. Cartier. A primer of Hopf algebras. In P. Cartier, B. Julia, P. Moussa, and P. Vanhove, editors, *Frontiers in number theory, physics, and geometry*, volume II, pages 537–615. Springer, 2007.

[14] A. Cayley. On the theory of the analytical forms called trees. *Philosophical Magazine Series 4*, 13(85), 1857.

[15] E. Celledoni, A. Marthinsen, and B. Owren. Commutator-free Lie group methods. *Future Generation Computer Systems*, 19(3):341–352, 2003.

[16] E. Celledoni, R.I. McLachlan, B. Owren, and G.R.W. Quispel. Energy-preserving integrators and the structure of B-series. *Foundations of Computational Mathematics*, 10:673–693, 2010.

[17] F. Chapoton. A rooted-trees q-series lifting a one-parameter family of Lie idempotents. *Algebra & Number Theory*, 3(6):611–636, 2009.

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

[19] P. Chartier, E. Hairer, and G. Vilmart. A substitution law for B-series vector fields. *INRIA report*, (5498), 2005.

[20] P. Chartier, E. Hairer, and G. Vilmart. Numerical integrators based on modified differential equations. *Mathematics of Computation*, 76(260):1941, 2007.

[21] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Communications in Mathematical Physics*, 199(1):203–242, 1998.

[22] P.E. Crouch and R. Grossman. Numerical integration of ordinary differential equations on manifolds. *Journal of Nonlinear Science*, 3(1):1–33, 1993.

[23] A. Dzhumadil'daev and C. Löfwall. Trees, free right-symmetric algebras, free Novikov algebras and identities. *Homology, Homotopy and applications*, 4(2):165–190, 2002.

[24] K. Ebrahimi-Fard, A. Lundervold, D. Manchon, H. Munthe-Kaas, and J.E. Vatne. On the post-Lie operad. *Preprint*, 2011.

[25] K. Ebrahimi-Fard and D. Manchon. A Magnus-and Fer-type formula in dendriform algebras. *Foundations of Computational Mathematics*, 9:1–22, 2009.

[26] K. Ebrahimi-Fard and D. Manchon. Pre-Lie Butcher series. *Preprint*, 2011.

[27] K. Engø. On the construction of geometric integrators in the RKMK class. *BIT Numerical Mathematics*, 40(1):41–61, 2000.

[28] K. Engø and A. Marthinsen. Modeling and solution of some mechanical problems on Lie groups. *Multibody System Dynamics*, 2(1):71–88, 1998.

[29] I. Gelfand, S. Gelfand, V. Retakh, and R.L. Wilson. Quasideterminants. *Advances in Mathematics*, 193(1):56–141, 2005.

[30] I.M. Gelfand and V.S. Retakh. Determinants of matrices over noncommutative rings. *Functional Analysis and Its Applications*, 25(2):91–102, 1991.

[31] M. Gerstenhaber. The cohomology structure of an associative ring. *Annals of Mathematics*, 78(2):267–288, 1963.

[32] R. Grossman and R.G. Larson. Hopf-algebraic structure of families of trees. *J. Algebra*, 126(1):184–210, 1989.

[33] E. Hairer. Backward analysis of numerical integrators and symplectic methods. *Annals of Numerical Mathematics*, 1(1-4):107–132, 1994.

[34] E. Hairer. Important aspects of geometric numerical integration. *Journal of Scientific Computing*, 25(1):67–81, 2005.

[35] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer, second edition, 2006.

[36] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving ordinary differential equations I: Nonstiff problems*. Springer, 1993.

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

[38] M. E. Hoffman. Combinatorics of rooted trees and hopf algebras. *Transactions of the American Mathematical Society*, 355(9):3795–3812, 2003.

[39] A. Iserles. Numerical methods on (and off) manifolds. In F. Cucker and M. Shub, editors, *Foundations of Computational Mathematics*, pages 180–189, 1997.

[40] A. Iserles, H.Z. Munthe-Kaas, S.P. Nørsett, and A. Zanna. Lie-group methods. *Acta numerica*, 9:215–365, 2000.

[41] A. Iserles, G.R.W. Quispel, and P.S.P. Tse. B-series methods cannot be volume-preserving. *BIT Numerical Mathematics*, 47(2):351–378, 2007.

[42] A. Iserles and A. Zanna. Qualitative numerical analysis of ordinary differential equations. In J. Renegar, M. Shub, and S. Smale, editors, *The mathematics of numerical analysis: 1995 AMS-SIAM Summer Seminar in Applied*

Mathematics, Utah, volume 32, page 421. American Mathematical Society, 1996.

[43] D. Kreimer. On the Hopf algebra structure of perturbative quantum field theories. *Advances in Theoretical and Mathematical Physics*, 2(2):303–334, 1998.

[44] D. Kreimer. Chen's iterated integral represents the operator product expansion. *Advances in Theoretical and Mathematical Physics*, 3(3):627–670, 1999.

[45] B. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*. Cambridge Univ Press, 2004.

[46] D. Lewis and P.J. Olver. Geometric integration algorithms on homogeneous manifolds. *Foundations of Computational Mathematics*, 2(4):363–392, 2002.

[47] A. Lundervold and H. Z. Munthe-Kaas. Hopf algebras of formal diffeomorphisms and numerical integration on manifolds. *Contemporary Mathematics*, 539:295–324, 2011.

[48] A. Lundervold and H.Z. Munthe-Kaas. On pre-Lie-type algebras with torsion and curvature. *Preprint*, 2010.

[49] A. Lundervold and H.Z. Munthe-Kaas. Backward error analysis and the substitution law for Lie group integrators. *Submitted*, 2011. ArXiv preprint math:1106.1071.

[50] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on pure and applied mathematics*, 7(4):649–673, 1954.

[51] D. Manchon. Hopf Algebras in Renormalisation. In M. Hazewinkel, editor, *Handbook of Algebra*, volume 5, pages 365–427. North Holland, 2008.

[52] D. Manchon. A short survey on pre-Lie algebras. Available at `http://math.univ-bpclermont.fr/~manchon/biblio/ESI-prelie2009.pdf`, 2009.

[53] R.I. McLachlan and G.R.W. Quispel. Six lectures on the geometric integration of ODEs. In R.A. DeVore, A. Iserles, and S. Endre, editors, *London Math. Soc. Lecture Note Series*, volume 284, pages 155–210. Cambridge Univ. Press, 2001.

[54] R.I. McLachlan and G.R.W. Quispel. Geometric integrators for ODEs. *Journal of Physics A: Mathematical and General*, 39:5251, 2006.

[55] H. Munthe-Kaas. Lie–Butcher theory for Runge–Kutta methods. *BIT Numerical Mathematics*, 35(4):572–587, 1995.

[56] H. Munthe-Kaas. Runge–Kutta methods on Lie groups. *BIT Numerical Mathematics*, 38(1):92–111, 1998.

[57] H. Munthe-Kaas. High order Runge-Kutta methods on manifolds. *Applied Numerical Mathematics*, 29(1):115–127, 1999.

[58] H. Munthe-Kaas and W. Wright. On the Hopf algebraic structure of Lie group integrators. *Foundations of Computational Mathematics*, 8(2):227–257, 2008.

[59] H. Munthe-Kaas and A. Zanna. Numerical integration of differential equations on homogeneous manifolds. In F. Cucker and M. Shub, editors, *Foundations of Computational Mathematics*, 1997.

[60] B. Owren. Order conditions for commutator-free Lie group methods. *Journal of Physics A: Mathematical and General*, 39, 2006.

[61] B. Owren and A. Marthinsen. Runge–Kutta methods adapted to manifolds and based on rigid frames. *BIT Numerical Mathematics*, 39(1):116–142, 1999.

[62] R.S. Palais. A global formulation of the Lie theory of transformation groups. *Memoirs of the AMS*, 22, 1957.

[63] C. Reutenauer. *Free Lie algebras*. Oxford University Press, 1993.

[64] J.M. Sanz-Serna and M.P. Calvo. *Numerical Hamiltonian problems*. Chapman & Hall/CRC, 1994.

[65] D. Segal. Free Left-Symmetrical Algebras and an Analogue of the Poincaré-Birkhoff-Witt Theorem. *Journal of Algebra*, 164(3):750–772, 1994.

[66] R.W. Sharpe. *Differential geometry: Cartan's generalization of Klein's Erlangen program*. Springer, 1997.

[67] M. Spivak. *A Comprehensive Introduction to Differential Geometry*, volume 1. Publish or Perish, third edition, 2005.

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

[69] P.S.P. Tse. *Geometric Numerical Integration: On the Numerical Preservation of Multiple Geometric Properties for Ordinary Differential Equations*. PhD thesis, La Trobe University, 2007.

[70] B. Vallette. Homology of generalized partition posets. *Journal of Pure and Applied Algebra*, 208(2):699–725, 2007.

[71] G. Vilmart. *Étude d'intégrateurs géométriques pour des équations différentielles*. PhD thesis, Université de Genève, 2008.

[72] E.B. Vinberg. Convex homogeneous cones. *Transactions of the Moscow Mathematical Society*, 12:340–403, 1963.

[73] R.F. Warming and B.J. Hyett. The modified equation approach to the stability and accuracy analysis of finite-difference methods. *Journal of Computational Physics*, 14(2):159–179, 1974.

[74] J.H. Wilkinson. Error analysis of floating-point computation. *Numerische Mathematik*, 2(1):319–340, 1960.

[75] J. Wisdom and M. Holman. Symplectic maps for the N-body problem. *The Astronomical Journal*, 102:1528–1538, 1991.