

B-series methods are exactly the affine equivariant methods

Robert I. McLachlan¹ · Klas Modin² ·
Hans Munthe-Kaas³ · Olivier Verdier⁴

Received: 2 September 2014 / Revised: 27 April 2015 / Published online: 4 August 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Butcher series, also called B-series, are a type of expansion, fundamental in the analysis of numerical integration. Numerical methods that can be expanded in B-series are defined in all dimensions, so they correspond to *sequences of maps*—one map for each dimension. A long-standing problem has been to characterise those sequences of maps that arise from B-series. This problem is solved here: we prove that a sequence of smooth maps between vector fields on affine spaces has a B-series expansion if and only if it is *affine equivariant*, meaning it respects all affine maps between affine spaces.

Mathematics Subject Classification 37C80 · 37C10 · 41A58

✉ Olivier Verdier
olivier.verdier@hib.no

Robert I. McLachlan
r.mclachlan@massey.ac.nz

Klas Modin
klas.modin@chalmers.se

Hans Munthe-Kaas
hans.munthe-kaas@math.uib.no

¹ Department of Mathematical Sciences, Massey University, Palmerston North, New Zealand

² Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, Göteborg, Sweden

³ Department of Mathematics, University of Bergen, Bergen, Norway

⁴ Department of Mathematics, Bergen University College, Bergen, Norway

1 Introduction

Let $\Phi(h, f): \mathbf{R}^n \rightarrow \mathbf{R}^n$ be a numerical time-stepping method for the differential equation

$$\dot{y} = f(y), \quad y(0) = y_0 \in \mathbf{R}^n.$$

That is, the time-stepping map $y_k \mapsto y_{k+1}$ is given by $y_{k+1} = \Phi(h, f)(y_k)$, where $y_k \approx y(hk)$. The convergence order of the method is obtained by comparing the Taylor expansion of $h \mapsto \Phi(h, f)(y_0)$ with the Taylor expansion of $h \mapsto y(h)$, using $\dot{y} = f(y)$ successively to avoid derivatives of y . If $\Phi(h, f)$ is a Runge–Kutta method, then the expansion is a linear combination of *elementary differentials* of f . For example, the first two terms for the midpoint method $y_{k+1} - y_k = hf \left(\frac{y_{k+1} + y_k}{2} \right)$ are

$$y_{k+1} - y_k = hf(y_k) + \frac{h^2}{2} f'(y_k)f(y_k) + \mathcal{O}(h^3).$$

To work out higher order terms by direct methods is tedious and results in long, convoluted tables (see, for example [18]).

In 1957, however, Merson [23] rediscovered a remarkable structure, found already by Cayley [8] in 1857: a one-to-one correspondence between elementary differentials and *rooted trees*. This structure is the basis of the influential work by Butcher, who, in 1963, gave the first modern treatment of the order conditions for Runge–Kutta methods [4], and, in 1974, developed an algebraic theory for series expansions of integration methods [5].

Let T denote the set of rooted trees. The expansion of a Runge–Kutta method is of the form

$$y_{k+1} - y_k = \sum_{\tau \in T} h^{|\tau|} \alpha(\tau) \mathcal{F}(\tau)[f](y_k), \tag{1}$$

where $|\tau|$ denotes the number of vertices of τ , $\alpha: T \rightarrow \mathbf{R}$ characterises the method, and $\mathcal{F}(\tau)[f]$ is the elementary differential of f associated with τ (see Sect. 3.6 for details). The right-hand side of (1) is called a *Butcher series*, or *B-series*, named so in 1974 by Hairer and Wanner [15]. The rich algebraic structure of B-series has since been studied extensively [7, 10, 11, 17, 25]. A numerical integration method $\Phi(h, f)$ whose expansion in h is of the form (1) is called a *B-series method*. In addition to numerical contexts, B-series have arisen in other branches of mathematics, such as noncommutative geometry, in models of renormalization [2, 3, 13] and rough paths theory [14].

Runge–Kutta methods are dense in the space of all B-series [6, § 317]: given any series of the form (1) and any $p \in \mathbf{N}$, there exists a Runge–Kutta method whose B-series coincides up to order h^p . There are, however, methods $\Phi(h, f)$ other than Runge–Kutta whose expansions in h are B-series. Examples are Rosenbrock methods like $y_{n+1} = y_n + h(I - \frac{1}{2}hf'(y_n))^{-1}f(y_n)$ [16], exponential integrators like $y_{n+1} = y_n + h\varphi(hf'(y_n))f(y_n)$ where $\varphi(z) = (e^z - 1)/z$, and the average vector field method $y_{n+1} = y_n + h \int_0^1 f(\xi y_{n+1} + (1 - \xi)y_n) d\xi$ [26].

So, which integration methods are B-series methods? Of course, given some method $\Phi(h, f)$, one can always check (1) by an expansion in h . But which properties *characterise* B-series methods? This is a natural, long-standing question that we answer

here. Our result is primarily based on two previous results: (i) that Runge–Kutta methods (and hence B-series methods) are *equivariant* with respect to affine transformations [21], and (ii) that local, equivariant maps can be expanded in a type of series described by *aromatic trees* [24]. Our result states that an integration method is a B-series method if and only if it defines an *affine equivariant sequence*, meaning chiefly that it is equivariant and keeps decoupled systems decoupled.

Before going into the details of the result, we explain a few key points, fundamental throughout the paper.

- Many ODE integration methods (in particular B-series methods) fulfill

$$\Phi(\varepsilon h, f) = \Phi(h, \varepsilon f)$$

for any positive ε . We may therefore disregard the dependency on h and write $\Phi(h, f) = \Phi(hf)$.

- We regard an ODE integration method as a map Φ from a neighbourhood of zero of the smooth, compactly supported vector fields $\mathfrak{X}_0(\mathbf{R}^n)$ to the set of diffeomorphisms $\text{Diff}(\mathbf{R}^n)$ (see Definition 2.1). To consider a neighbourhood of zero means to restrict the integration method to sufficiently small time steps. Then Φ is an approximation of the *exponential map* $\exp: \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \text{Diff}(\mathbf{R}^n)$.
- We take the *backward error analysis* point-of-view, which represents an integration method $\Phi(f)$ as $\Phi = \exp \circ \phi$ for some map $\phi: \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$. Here, equality holds in the sense of formal power series. To avoid technical questions of convergence subordinate to our aim, the main result, Theorem 2.4, is formulated for maps from vector fields to vector fields, independently of their relation to numerical integration methods. The same argument is used in [24, § 2.2]. The key observation is, nevertheless, that Φ is a B-series method if and only if $\phi(f)$ can be expanded in a B-series, i.e.,

$$\phi(f) = \sum_{\tau \in T} \beta(\tau) \mathcal{F}(\tau)[f], \tag{2}$$

for some map $\beta: T \rightarrow \mathbf{R}$. Each term in (1) is a homogeneous polynomial in f , so each term corresponds to a symmetric, multi-linear map

$$\mathfrak{X}_0(\mathbf{R}^n) \times \cdots \times \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$$

evaluated at f . For instance, the term $f'f$ in a B-series corresponds to the bilinear map $(f, g) \mapsto (f'g + g'f)/2$. Consequently, (1) is the Taylor series of $\phi: \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$, so our investigation consists of classifying those maps ϕ whose Taylor series are B-series.

- An ODE integration method actually corresponds to a *sequence* of maps: one for each dimension $n \in \mathbf{N}$. From here on, we therefore use ϕ to denote a sequence of maps $\{\phi_n\}_{n \in \mathbf{N}}$, where $\phi_n: \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$. This point of view is essential in the characterisation of B-series maps (see Sect. 2 for details).

The paper is organised as follows. The main result is stated in Sect. 2. In Sect. 3 we give preliminary results necessary for the proof. The main part of the proof is contained in Sect. 5, and uses results from Sect. 4 on special vector fields. Finally, Sect. 6 connects the core result from Sect. 5 to the main result as stated in Sect. 2.

2 Main result

Our main result is a simple criterion to decide whether a method is a B-series method. The essence of the result is captured as follows.

Let $\Phi = \{\Phi_n\}_{n \in \mathbb{N}}$ be an integration method, defined for all vector fields on all dimensions n . Then Φ is a B-series method if and only if the property of *affine equivariance* is fulfilled: if $a(x) := Ax + b$ is an affine map from \mathbb{R}^m to \mathbb{R}^n , and $f \in \mathfrak{X}_0(\mathbb{R}^m)$, $g \in \mathfrak{X}_0(\mathbb{R}^n)$ fulfil $g(Ax + b) = Af(x)$, then $a \circ \Phi_m(f) = \Phi_n(g) \circ a$.

The rigorous version of this result, using, as explained, the backward error analysis point of view, is stated in Theorem 2.4 at the end of this section. Before that, we need to define the essential concept of equivariance.

Our definition of equivariance is an extended version of that in [24, § 2.4], [21, § 4.3]. We first define the main object of study: sequences of smooth maps from a neighbourhood of zero in the space of compactly supported vector fields to itself. As we reuse this object several times, we encapsulate it in the following definition.

Definition 2.1 (*Integrator map*) Let $\mathfrak{X}_0(\mathbb{R}^n)$ be the space of compactly supported vector fields on \mathbb{R}^n with the test function topology. We define an *integrator map* as a sequence of smooth maps

$$\phi = \{\phi_n : U_n \subset \mathfrak{X}_0(\mathbb{R}^n) \rightarrow \mathfrak{X}_0(\mathbb{R}^n)\}_{n \in \mathbb{N}},$$

where each U_n is a suitable neighbourhood of the zero vector field in $\mathfrak{X}_0(\mathbb{R}^n)$.

In order for an integrator map ϕ to correspond to a B-series, there must clearly be some relationship between the individual maps ϕ_n . We denote by $\text{aff}(\mathbb{R}^n, \mathbb{R}^m)$ the set of affine maps:

$$\text{aff}(\mathbb{R}^n, \mathbb{R}^m) = \{a : \mathbb{R}^n \rightarrow \mathbb{R}^m \mid a(x) = Ax + b, A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n \text{ and } b \in \mathbb{R}^m\}.$$

Pullback of vector fields along invertible diffeomorphisms is generalised for non-invertible maps by the concept of *intertwining* (relatedness) of vector fields. We say that $a \in \text{aff}(\mathbb{R}^n, \mathbb{R}^m)$ *intertwines* the vector fields $f \in \mathfrak{X}_0(\mathbb{R}^n)$ and $g \in \mathfrak{X}_0(\mathbb{R}^m)$, which we denote by $f \overset{a}{\rightsquigarrow} g$, if

$$g(Ax + b) = Af(x). \tag{3}$$

Definition 2.2 (*Affine equivariance*) An integrator map ϕ (Definition 2.1) is called *affine equivariant* if, for all $m, n \in \mathbf{N}$ and all $a \in \text{aff}(\mathbf{R}^n, \mathbf{R}^m)$,

$$f \overset{a}{\rightsquigarrow} g \implies \phi_m(f) \overset{a}{\rightsquigarrow} \phi_n(g).$$

Here are some of the properties attributed to affine equivariance:

1. Integrator maps that are equivariant with respect to *invertible* affine maps preserve affine symmetries. At every fixed dimension, the integration method corresponding to such an integrator map is affine equivariant in the standard sense, used in [24].
2. Consider systems of the form

$$\begin{aligned} \dot{x} &= f(x), \\ \dot{y} &= g(x, y) \end{aligned}$$

where $x \in \mathbf{R}^m, y \in \mathbf{R}^n$. Geometrically, they are characterised by preservation of the foliation by planes $x = \text{constant}$ [22]. Integrator maps ϕ (and corresponding integration methods) that are equivariant with respect to affine *surjections*, in this case the surjection $(x, y) \mapsto x$, preserve this foliation. In addition, they are “closed with respect to closed subsystems” [1]: the map $\phi_{m+n}((f, g))$ restricted to x in domain and range is identical to $\phi_m(f)$.

3. Integrator maps (and corresponding integration methods) that are equivariant with respect to affine *injections* preserve affine weak integrals. (Recall that a weak integral is a function $I : \mathbf{R}^n \rightarrow \mathbf{R}$ such that $I = 0$ implies $\dot{I} = 0$; an affine weak integral is equivalent to an invariant affine subspace.) In addition, the map on the affine subspace induced by such an integrator map is identical to that produced by application directly to the system on the subspace.

Each of these are also properties of B-series methods, while non-B-series methods, such as partitioned Runge–Kutta methods, do not have them. In addition, B-series methods have many other structural and geometric properties [12, 19]. Perhaps surprisingly, all of these are now seen to be consequences of affine equivariance.

To give a rigorous definition of integrator maps corresponding to B-series methods, we need Taylor series of vector field maps. Let $\phi_n : \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$ be smooth. Its k th derivative at $0 \in \mathfrak{X}_0(\mathbf{R}^n)$, denoted $D^k \phi_n(0)$, is a k -linear, symmetric form on $\mathfrak{X}_0(\mathbf{R}^n)$. Taylor’s formula [20, Theorem I.5.12] states that

$$\begin{aligned} \phi_n(f) &= \phi_n(0) + \frac{D\phi_n(0)[f]}{1!} + \dots + \frac{D^k \phi_n(0) \overbrace{[f, \dots, f]}^k}{k!} \\ &\quad + \int_0^1 \frac{(1 - \sigma)^k}{k!} D^{k+1} \phi(\sigma f) \overbrace{[f, \dots, f]}^{k+1} d\sigma. \end{aligned}$$

Let, as before, T denote the set of rooted trees and $\mathcal{F}(\tau)[f] \in \mathfrak{X}_0(\mathbf{R}^n)$ denote the elementary differential of $f \in \mathfrak{X}_0(\mathbf{R}^n)$ associated with $\tau \in T$. Further, let $\langle T_k \rangle$ denote the free \mathbf{R} -vector space over the set T_k of trees with k vertices. That is, each element in $\langle T_k \rangle$ is an \mathbf{R} -linear combination of elements in T_k . By construction, T_k is a basis for $\langle T_k \rangle$.

For each $k \in \mathbf{N}$ and $f \in \mathfrak{X}_0(\mathbf{R}^n)$, $\mathcal{F}(\cdot)[f]$ is naturally extended to a linear map $\langle T_k \rangle \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$. We define B-series maps as those integrator map whose Taylor coefficients are elementary differentials.

Definition 2.3 (*B-series map*) An integrator map ϕ (Definition 2.1) is called a *B-series map* if, for each $k \in \mathbf{N}$, there exists $\tau_k \in \langle T_k \rangle$ such that for all $n \in \mathbf{N}$

$$\frac{D^k \phi_n(0)[f, \dots, f]}{k!} = \mathcal{F}(\tau_k)[f], \quad \forall f \in \mathfrak{X}_0(\mathbf{R}^n).$$

We are now in a position to state the main result of this paper.

Theorem 2.4 *Let ϕ be an integrator map (Definition 2.1). Then ϕ is a B-series map (Definition 2.3) if and only if it is affine equivariant (Definition 2.2).*

Proof Using Propositions 6.2 and 6.3, this result is equivalent to Theorem 5.1. □

Remark 2.5 It turns out that, although B-series are affine equivariant, it is not necessary to use equivariance with respect to all affine maps to prove Theorem 2.4. We only use equivariance with respect to all *surjective* affine maps to prove Theorem 5.1 (invertible affine maps in Proposition 3.2 and surjective affine maps in Lemma 4.2) and with respect to the trivial affine injections from \mathbf{R}^0 to \mathbf{R}^n (in Lemma 6.1).

2.1 Idea of the proof

The proof of Theorem 2.4 is long and contains many details. In this section we therefore explain its main ingredients by proving Theorem 2.4 in the special case when the functions ϕ_m are *homogeneous quadratic* maps, instead of arbitrary nonlinear smooth maps.

We use a transfer argument, similar to that of [24], to transfer the statement from affine equivariant integrator maps, to linear equivariant polynomials P_m in the derivatives of the vector fields (this is Propositions 6.2 and 6.3).

After the transfer argument, we have for every dimension m a vector valued homogeneous quadratic polynomial P_m in the derivatives of the vector field at zero (cf. Sect. 3.2). A component of such a vector could be for $f_2^3 f_1^1$ in dimension three, or a linear combination of such expressions (we follow standard practice for the notation of components and partial derivatives, which are detailed in Sect. 3).

We make explicit the assumption that the sequence P_m is linear equivariant. Consider an arbitrary linear map A from \mathbf{R}^n to \mathbf{R}^m . We say that A intertwines a vector field f in dimension n and a vector field g in dimension m , if $g(Ax) = Af(x)$ for any x in \mathbf{R}^n (cf. Sect. 3.3). The sequence $\{P_m\}_{m \in \mathbf{N}}$ is equivariant if whenever A intertwines f and g , it also intertwines $P_n[f]$ and $P_m[g]$ (cf. Sect. 3.4).

We now prove the special case of quadratic linear equivariant functions.

Proposition 2.6 *Suppose that $P = \{P_m\}_{m \in \mathbf{N}}$ is a vector-valued, homogeneous quadratic polynomial in the derivatives of vector fields at zero. Suppose further that*

for any linear map $A: \mathbf{R}^n \rightarrow \mathbf{R}^m$ and vector fields such that $g(Ax) = Af(x)$, we have $P[g](Ax) = AP[f](x)$. Then P must be of the form

$$P_m[f] = \lambda \sum_{i,j=1}^m f_i^j f^i \partial_j$$

for some scalar λ .

We use the following standard notations (see Sect. 3.6 for the notations pertaining to elementary differentials used in this paper):

$$\begin{aligned} \mathfrak{I}[f] &= \sum_{i,j=1}^n f_j^i f^j \partial_i, \\ \mathfrak{O} \bullet [f] &= \sum_{i,j=1}^n f_i^j f^j \partial_j. \end{aligned} \tag{4}$$

The main ingredient of the proof is to notice that the term $\mathfrak{O} \bullet$ is coupling (cf. Lemmas 4.2 and 4.5), which contradicts equivariance (cf. Lemma 4.9).

Proof Consider an arbitrary vector field f in dimension k (in this proof, k will take the value 1 or 2). For any dimension $m \geq k$, we denote the vector $f \oplus 0$ as the vector equal to f on the first k components, padded with zeros (cf. Sect. 4.1), i.e.,

$$(f \oplus 0)(x_1, \dots, x_k, \dots, x_m) := (f^1(x_1, \dots, x_k), \dots, f^k(x_1, \dots, x_k), 0, \dots, 0).$$

For $m \geq k$, we also define the linear projection $\Pi_k^m: \mathbf{R}^m \rightarrow \mathbf{R}^k$ which keeps the first k components. Note that the projection Π_k^m intertwines $f \oplus 0$ and f , since $\Pi_k^m(f \oplus 0)(x) = f(\Pi_k^m x)$ (cf. Proposition 4.1). Finally, observe that from the formulas (4), we have for any scalars λ and μ (cf. Lemma 4.6)

$$\Pi_k^m(\lambda \mathfrak{I} + \mu \mathfrak{O} \bullet)[f \oplus 0] = (\lambda \mathfrak{I} + \mu \mathfrak{O} \bullet)[f]. \tag{5}$$

Given these definitions and observations, the proof proceeds as follows.

- Using equivariance with respect to invertible linear maps on each dimension, we obtain from [24] (see Proposition 3.2) that, for every dimension m , we have

$$P_m[f] = (\lambda_m \mathfrak{I} + \mu_m \mathfrak{O} \bullet)[f]$$

for some real numbers λ_m and μ_m .

- We now proceed to show that $P_m[f] = (\lambda_2 \mathfrak{I} + \mu_2 \mathfrak{O} \bullet)[f]$ for any vector field f in m dimensions (cf. Lemma 5.7). We first show that for $m \geq 2$ we have $\lambda_m = \lambda_2$ and $\mu_m = \mu_2$, and then show that, on one-dimensional vector fields, $P_1[f] = (\lambda_2 \mathfrak{I} + \mu_2 \mathfrak{O} \bullet)[f]$.

- (a) Consider a vector field f in dimension two. As Π_2^m intertwines $f \oplus 0$ and f , equivariance of P leads to $\Pi_2^m P_m[f \oplus 0](x) = P_2[f](\Pi_2^m x) = (\lambda_2 \mathfrak{I} + \mu_2 \mathfrak{O} \bullet)[f](\Pi_2^m x)$.

Combining this with (5), we obtain that $\lambda_m \mathfrak{I} + \mu_m \mathfrak{O} \bullet$ and $\lambda_2 \mathfrak{I} + \mu_2 \mathfrak{O} \bullet$ coincide on vector fields in dimension two. Appealing again to [24] (see Proposition 3.2), we obtain that $\lambda_m = \lambda_2$ and $\mu_m = \mu_2$.

- (b) Consider now a vector field f in one dimension. Construct $f \oplus 0$ in two dimensions; equivariance now gives $\Pi_1^2 P_2[f \oplus 0](x) = P_1[f](\Pi_1^2 x)$. Again, by comparing this with (2.1), we obtain that $P_1[f] = (\lambda_2 \mathfrak{I} + \mu_2 \mathfrak{O} \bullet)[f]$.

Combining the last two items, we deduce that in fact $P_m[f] = (\lambda_2 \mathfrak{I} + \mu_2 \mathfrak{O} \bullet)[f]$ for any integer m .

3. The final step is now to show that $\mu_2 = 0$ (cf. Lemma 5.8). The idea is to apply P_2 to the special vector field $g(x_1, x_2) = (1, x_2)$ (cf. Lemma 4.10 for the general expression of such special vector fields). One checks that, at $x = 0$, we have $\mathfrak{I}[g](0) = 0$, and that $\mathfrak{O} \bullet[g](0) = (1, 0)$. We conclude that $P_2[f](0) = \mu_2(1, 0)$, and in particular $\Pi_1^2 P_2[f](0) = \mu_2$. Denote by 1 the constant vector field on \mathbf{R} ; it is clear that Π_1^2 intertwines g and 1 . The equivariance of P now implies that $\Pi_1^2 P_2[g](0) = P_1[1](0) = 0$, which entails $\mu_2 = 0$.

We conclude that for any integer m , $P_m = \lambda_2 \mathfrak{I}$, which, recalling the notation (4), is the claim of the proposition. □

The rest of the paper consists of generalizing the arguments in the proof above, in order to accommodate homogeneous polynomials of any degree.

3 Preliminary definitions

3.1 Polynomial vector fields

For a fixed dimension n , we define an infinite dimensional vector space F^n of *polynomials* of arbitrary degree. We use derivatives as coordinates in that space. These coordinates are thus indexed by the partial derivatives, as

$$(f, f_1, \dots, f_n, f_{11}, f_{12}, \dots)$$

with appropriate symmetry conditions, such as $f_{12} = f_{21}$.

We denote by X^n the set of *vector-valued polynomials*, which consists of n elements of F^n , that is $X^n = (F^n)^n$. An element in X^n should be regarded as a *polynomial vector field*.

3.2 Forms

A *k-form* in dimension n is a homogeneous polynomial of degree k on the space of polynomial vector fields X^n . We denote scalar k -forms on X^n by $S^k(X^n)$. A *vector*

valued k -form in dimension n is a list of n k -forms, regarded as a vector. It is thus an element of $\mathbf{R}^n \otimes S^k(X^n)$. As is customary, we use the basis ∂_i . Note that we use the same notation for the basis in all dimensions, which should not cause confusion. For instance, in dimension n the map η defined by

$$\eta[f] = f^1 f_1^1 \partial_1 + f_1^n f_n^n \partial_n$$

is a vector-valued 2-form. In one dimension, the coordinates are $(f^1, f_1^1, f_{11}^1, \dots)$, corresponding to (f, f', f'', \dots) , and a 2-form in one dimension could be $ff' + (f'')^2$. In two dimensions, an example of a vector valued 3-form is $P[f] = f_{112}^2 f^2 f^1 \partial_1 + (f_2^1)^3 \partial_2$.

3.3 Intertwining

Given a linear map $A \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ we say that A *intertwines* $f \in X^n$ with $g \in X^m$, denoted

$$f \overset{A}{\rightsquigarrow} g,$$

if the equality

$$g(Ax) = Af(x) \tag{6}$$

is valid for all $x \in \mathbf{R}^n$.

We give an example of intertwining with respect to a projection. Define the scalar polynomials $f_1 \in F^1$ and $f_2 \in F^2$, the polynomial vector field $f \in X^2$ by $f(x_1, x_2) := f_1(x_1)\partial_1 + f_2(x_1, x_2)\partial_2$, and $g \in X^1$ by $g(x) = f_1(x)\partial_1$. We denote the projection $\pi \in \mathcal{L}(\mathbf{R}^2, \mathbf{R}^1)$ on the first coordinate, that is, $\pi(x_1, x_2) = x_1$. In that case, one can check that $g(\pi(x)) = \pi(f(x))$, so we have $f \overset{\pi}{\rightsquigarrow} g$.

We now give an example of intertwining with respect to an injection. Define the vector field $f \in X^1$ by $f = f_1(x_1)\partial_1$, where $f_1 \in F^1$. Consider now the vector field $g \in X^2$ defined by $g(x_1, x_2) = g_1(x_1, x_2)\partial_1 + g_2(x_1, x_2)\partial_2$, with the property that $g_2(x, 0) = 0$ and $g_1(x, 0) = f_1(x)$ for any $x \in \mathbf{R}$. Define the injection $i \in \mathcal{L}(\mathbf{R}, \mathbf{R}^2)$ by $i(x) = (x, 0)$. As $g(i(x)) = i(f(x))$, we have $f \overset{i}{\rightsquigarrow} g$.

3.4 Equivariant sequences

We define an *equivariant sequence of (vector-valued) k -forms* as a sequence η_n of vector-valued k -forms with the property that whenever

$$f \overset{A}{\rightsquigarrow} g, \tag{7}$$

it holds that

$$\eta_n(f) \overset{A}{\rightsquigarrow} \eta_m(g).$$

A typical example of an equivariant sequence of 2-forms is

$$\eta_n[f] := \sum_{i,k=1}^n f_i^k f^i \partial_k.$$

Finally, we use the following simplifying notation. If $f \in X^n$ and P is a sequence of k -forms, we use the notation

$$P[f] := P_n[f].$$

If the dimension n is not clear from the context, we use the explicit form $P_n[f]$.

3.5 Aromatic forests, trees and molecules

We now review some definitions from [24]. Let Γ denote the set of all directed graphs with a finite number of vertices, where each vertex has zero or one outgoing edges. A vertex with no outgoing edges is called a *root*. For $\gamma \in \Gamma$, let $\mathcal{V}(\gamma)$ and $\mathcal{E}(\gamma)$ denote the vertices and edges of the graph, let $\mathcal{R}(\gamma) \subset \mathcal{V}(\gamma)$ denote the *root vertices*. For $v \in \mathcal{V}(\gamma)$, let $\mathcal{P}(v) \subset \mathcal{V}(\gamma)$ denote the set of *parent vertices*, $\mathcal{P}(v) := \{p \in \mathcal{V}(\gamma) \mid (p, v) \in \mathcal{E}(\gamma)\}$. Let $|\gamma| := \#\mathcal{V}(\gamma)$ denote the number of vertices and $|\mathcal{R}(\gamma)|$ the number of roots.

We have $\Gamma = \bigcup_{r=0, k=1}^{\infty} \Gamma_k^r$, where Γ_k^r denotes graphs with r roots and k vertices. We denote $\Gamma^r := \bigcup_{k=1}^{\infty} \Gamma_k^r$. Let $\langle \Gamma \rangle$ and $\langle \Gamma^r \rangle$ denote the free \mathbf{R} -vector spaces over Γ and Γ^r . An element of Γ^1 is called an *aromatic tree*. For instance, the following is an element of Γ^1 , as it has one root, so it is an aromatic tree:



By convention, we will assume that the cycles are always oriented counterclockwise and we will draw the aromatic trees in short form as:



The set of *trees* is the subset $T \subset \Gamma^1$ of connected graphs in Γ^1 . Similarly, the set of *aromatic molecules* is the subset $M \subset \Gamma^0$ of connected graphs in Γ^0 .

We define the *product of graphs* as their disjoint union: for $\gamma_1, \gamma_2 \in \Gamma$ the product $\gamma_1 \gamma_2 = \gamma_2 \gamma_1$ is the graph consisting of the union of the vertices and edges of the two graphs.

Lemma 3.1 *Let $\gamma \in \Gamma^1$. Then γ can be decomposed as*

$$\gamma = \mu_1 \mu_2 \dots \mu_k \tau,$$

where $\mu_1, \dots, \mu_k \in M$ and $\tau \in T$.

Proof Any graph $\gamma \in \Gamma^1$ can be decomposed into a union of its connected components, where each connected component is either in T or in M . As γ has one root, the root must belong to one of the components, which is thus in T . The other components are also aromatic forests, but with the same number of nodes and arrows, so they cannot have any root and must belong to M . \square

3.6 Elementary differentials

Consider $f \in X^n$ and $\gamma \in \Gamma_k^r$. The set $\mathcal{V}(\gamma)$ denotes the set of vertices of γ . For a node $i \in \mathcal{V}(\gamma)$, we denote by $\mathcal{P}(i)$ the parent vertices of i . We define the *elementary differential* $\mathcal{F}_n(\gamma)[f] \in (\mathbf{R}^n)^{\otimes r} \otimes S^k(X^n)$ in tensor component notation as

$$\mathcal{F}_n(\gamma)[f] := \prod_{i \in \mathcal{V}(\gamma)} f_{\mathcal{P}(i)}^i, \tag{8}$$

where we use the *Einstein summation convention*: repeated indexes are summed over in the range $\{1 \dots n\}$. Every lower index is paired with an upper index, but the upper indices that correspond to roots are not paired. We rewrite that definition in a more tractable form in (9). Note that we make sure to keep track of the *dimension* n in those expressions. Consider, for instance, the expression

$$\mathcal{F}_n(\bullet \circlearrowleft \bullet)[f] = \sum_{i=1}^n f^i \partial_i \sum_{j=1}^n f_j^j.$$

In one dimension it is

$$\mathcal{F}_1(\bullet \circlearrowleft \bullet)[f] = f^1 f_1^1 \partial_1 \equiv f f'.$$

In two dimensions it is

$$\mathcal{F}_2(\bullet \circlearrowleft \bullet)[f] = (f^1 \partial_1 + f^2 \partial_2)(f_1^1 + f_2^2) = f \operatorname{div} f.$$

Below is an example of an elementary differential where $\gamma \in M$ and $\mathcal{F}_n(\gamma)[f]$ is a scalar two-form:

$$\mathcal{F}_n(\bullet \circlearrowright \bullet)[f] = f^i f_{ij}^j = \sum_{i=1}^n \sum_{j=1}^n f^i \frac{\partial^2 f^j}{\partial x_i \partial x_j} = f \cdot \operatorname{grad}(\operatorname{div} f).$$

Here, i is the the top vertex of γ , giving the part f^i (no parents, hence no subscript), and j is the bottom vertex of γ , giving the factor f_{ij}^j , since $\mathcal{P}(j) = \{i, j\}$.

We now rewrite the definition (8) of the elementary differential $\mathcal{F}_n(\gamma)[\cdot]: X^n \rightarrow X^n$ in an equivalent, but, for our purpose, more tractable form:

$$\mathcal{F}_n(\gamma)[f] = \sum_{v: \mathcal{V}(\gamma) \rightarrow [n]} \prod_{v \in \mathcal{V}(\gamma)} f_{\mathcal{P}(v)}^{v(v)} \prod_{r \in \mathcal{R}(\gamma)} \partial_{v(r)}, \tag{9}$$

Here, $[n] := \{1, 2, \dots, n\}$, $\mathcal{R}(\gamma)$ denotes the root vertices of the graph γ , $\partial_j \in \mathbf{R}^n$ denotes the unit vector in a direction $j \in [n]$, and $\mathcal{P}(v)$ denotes the set of parent vertices of vertex v , i.e., $\mathcal{P}(v) = \{p \in \mathcal{V}(\gamma) | (p, v) \in \mathcal{E}(\gamma)\}$. The sum runs over all possible maps $v: \mathcal{V}(\gamma) \rightarrow [n]$, assigning vertices in γ to integers in $[n]$.

To verify the re-writing (9), we note that an assignment of vertices in γ to integers in $[n]$ is already implicit in our interpretation of (8) where i both denotes a node in γ and an integer in $[n]$. Thus f_j^i in (8) is the same as $f_{\nu(\mathcal{P}(v))}^{\nu(v)}$ in (9). In (8), the root nodes are left as tensor components which are not summed over. In (9) we pair the root nodes $\nu(r)$ with the basis vector field $\partial_{\nu(r)}$, and hence the sum here runs over all possible maps $\nu: \mathcal{V}(\gamma) \rightarrow [n]$.

For each dimension n , Eq. (9) defines the *elementary differential map* \mathcal{F}_n . In particular, on the subspace $\langle \Gamma_k^1 \rangle$ of aromatic trees with k vertices, we have

$$\mathcal{F}_n: \langle \Gamma_k^1 \rangle \rightarrow \mathbf{R}^n \otimes S^k(X^n).$$

For fixed dimension n , the map \mathcal{F}_n is not injective. For instance,

$$\mathcal{F}_2 \left(\bullet \circ \circ + 2 \begin{array}{c} \bullet \\ | \\ \bullet \end{array} - 2 \begin{array}{c} \bullet \\ | \\ \bullet \end{array} \circ - \bullet \circ \circ \right) [f] = 0 \quad \text{for all } f \in X^2.$$

The elementary differential map is not injective even when restricted to $\langle T \rangle$. For instance,

$$\mathcal{F}_1 \left(\begin{array}{c} \bullet \\ | \\ \bullet \\ | \\ \bullet \end{array} - \begin{array}{c} \bullet \\ | \\ \bullet \\ | \\ \bullet \end{array} \right) [f] = 0 \quad \text{for all } f \in X^1.$$

This is one of the motivations for regarding the elementary differential as acting on all dimensions. Indeed, we build a *sequence*

$$\mathcal{F}: \langle \Gamma_k^1 \rangle \rightarrow \{\text{sequences of vector-valued } k\text{-forms}\}$$

defined by

$$\mathcal{F}(\gamma) := (\mathcal{F}_1(\gamma), \mathcal{F}_2(\gamma), \dots).$$

In the sequel, we use the following simplified notation: for $\gamma \in \Gamma$ and $f \in X^n$, we define

$$\gamma[f] := \mathcal{F}_n(\gamma)[f].$$

Note that the dimension n is implicitly defined by the space X^n that f belongs to. When ambiguity remains, we resume the explicit notation $\mathcal{F}_n(\gamma)[f]$.

As a first result for the elementary differential map, consider the following result, established in [24, § 7.4]:

Proposition 3.2 *For each dimension $n \in \mathbf{N}$ and degree $k \in \mathbf{N}$, \mathcal{F}_n is a surjection from $\langle \Gamma_k^1 \rangle$ to the space $\mathbf{R}^n \otimes S^k(X^n)$ of vector-valued equivariant k -forms in dimension n . Moreover, if $k \leq n$, then \mathcal{F}_n is a bijection.*

4 Special vector fields

The proof of Theorem 5.1 below is based on the construction of special vector fields. In particular, we need vector fields with block-diagonal Jacobians. We also need special vector fields that form a dual basis with respect to aromatic trees and molecules.

4.1 Partitioned vector fields

Vector fields with block-diagonal Jacobian matrices are called *partitioned*. They serve a special role in the sequel.

Proposition 4.1 *Consider the decomposition*

$$\mathbf{R}^{n+m} = \mathbf{R}^n \oplus \mathbf{R}^m,$$

and denote the associated projections by π_1 and π_2 . Then, given $f \in X^n$ and $g \in X^m$, there is a unique vector field $h \in X^{n+m}$ characterized by

$$h \overset{\pi_1}{\rightsquigarrow} f, \quad h \overset{\pi_2}{\rightsquigarrow} g.$$

That vector field is denoted

$$f \oplus g := h.$$

Proof $h \overset{\pi_1}{\rightsquigarrow} f$ means by definition (6) that $\pi_1 h(x, y) = f(x)$, where we denote a point $(x, y) \in \mathbf{R}^{n+m}$ so that $\pi_1(x, y) = x$. This means that $h(x, y)$ has $f(x)$ as first components, and likewise, $g(y)$ as last components. \square

We thus immediately obtain the following property of equivariant sequences.

Lemma 4.2 *If $f \in X^n$ and $g \in X^m$, and P is an equivariant sequence of k -forms, we have*

$$P[f \oplus g] = P[f] \oplus P[g].$$

Proof Consider the vector field $h = f \oplus g$, defined as in Proposition 4.1. By definition of the equivariance of the sequence (7), we have $P[h] \overset{\pi_1}{\rightsquigarrow} P[f]$, and $P[h] \overset{\pi_2}{\rightsquigarrow} P[g]$, which, again by Proposition 4.1 ensures that $P[h] = P[f] \oplus P[g]$. \square

Thus, equivariant sequences “keep decoupled systems decoupled”; this is essentially the difference between aromatic series and B-series.

We now derive special formulas for elementary differentials of partitioned vector fields. To do that, we first reformulate the elementary differential formula (9) using *dependency graphs*.

Definition 4.3 The *dependency graph* of $f \in X^n$, denoted $\text{dep}(f)$, is the directed, labeled graph defined by

$$\begin{aligned} \mathcal{V}(\text{dep}(f)) &= [n] \\ (j, i) \notin \mathcal{E}(\text{dep}(f)) &\iff \partial_j f^i(x) = 0 \quad \forall x \in \mathbf{R}^n. \end{aligned}$$

Note that a vector field is partitioned if and only if its dependency graph is disconnected.

Lemma 4.4 For $\mu \in M$ the elementary differential is given as

$$\mu[f] = \sum_{v \in \text{hom}(\mu, \text{dep}(f))} \prod_{v \in \mathcal{V}(\mu)} f_v^{\nu(v)}, \tag{10}$$

where $\text{hom}(\mu, \text{dep}(f))$ denotes graph homomorphisms of μ into $\text{dep}(f)$, i.e. a map of graphs sending vertices to vertices and edges to edges. For $\tau \in T$, component k of the elementary differential is given as

$$\tau[f]^k = \sum_{\substack{v \in \text{hom}(\tau, \text{dep}(f)) \\ v(\mathcal{R}(\tau))=k}} \prod_{v \in \mathcal{V}(\mu)} f_v^{\nu(v)}.$$

Proof Equation (9) expresses the elementary differential as a sum over all possible maps $v: \mathcal{V}(\gamma) \rightarrow \mathcal{V}(\text{dep}(f))$. By definition of the dependency graph Definition 4.3, we see that all maps which are not sending edges to edges must yield 0. Hence, we can restrict the sum to homomorphisms. Both formulas follow from this argument; in the first case μ has no roots. In the latter case component k is the multiplier in front of $\partial_{v(\mathcal{R}(\tau))}$, i.e., $k = v(\mathcal{R}(\tau))$. Thus we restrict to all homomorphisms sending the root of τ to k . □

Our next result shows that trees and molecules preserve, in a sense, the structure of partitioned vector fields.

Lemma 4.5 Consider the partitioned vector field $f = f_1 \oplus f_2 \in X^{m+n}$, with $f_1 \in X^m$ and $f_2 \in X^n$. If $\mu \in M$ and $\tau \in T$, then

$$\mu[f_1 \oplus f_2] = \mu[f_1] + \mu[f_2] \tag{11}$$

$$\tau[f_1 \oplus f_2] = \tau[f_1] \oplus \tau[f_2]. \tag{12}$$

Proof The dependencies $\delta = \text{dep } f$ decomposes $\delta = \delta_1 \delta_2$ in two disjoint graphs $\delta_1 = \text{dep}(f_1)$, $\delta_2 = \text{dep}(f_2)$. Since μ is connected, $\text{hom}(\mu, \delta) = \text{hom}(\mu, \delta_1) \cup \text{hom}(\mu, \delta_2)$ and the sum in Lemma 4.4 splits accordingly, yielding (Lemma 4.5). Similarly $\text{hom}(\tau, \delta) = \text{hom}(\tau, \delta_1) \cup \text{hom}(\tau, \delta_2)$ yields (Lemma 4.5). □

As opposed to trees and molecules, aromatic trees (which by Lemma 3.1 are products of molecules and a tree) do not preserve the structure of partitioned vector fields. This is the key to the characterisation of B-series. In the special case, however, when

the partition represents an injection of a vector field in a higher dimensional space, aromatic trees do preserve the partitioned structure.

Lemma 4.6 *If $f = g \oplus 0$ with $f \in X^n$ and $g \in X^m$, then*

$$\gamma[g \oplus 0] = \gamma[g] \oplus 0 \quad \forall \gamma \in \langle \Gamma_k^1 \rangle.$$

Proof The result follows from the elementary differential formula (9). We have to prove that the term corresponding to $v : \mathcal{V}(\gamma) \rightarrow [n]$ is zero whenever $v(\mathcal{V}(\gamma)) \not\subseteq [m]$. This is clear since there is then a vertex v of γ such that $v(v) \in [n] \setminus [m]$, and we use that $f_j^{v(v)} = 0$ for any derivative J . □

4.2 Dual vector fields

In classical B-series theory, results on linear independence of elementary differentials are obtained by specially constructed vector fields. We use the same technique as in [5, 19] to construct such vector fields.

First, we need to define the symmetry of a graph. Let $\sigma(\gamma)$ denotes the number of symmetries of a graph, defined as the size of the automorphism group of the graph, i.e.,

$$\sigma(\gamma) := \#\text{Aut}(\gamma)$$

where

$$\text{Aut}(\gamma) := \{v \in \text{Aut}(\mathcal{V}(\gamma)) \mid v(\mathcal{E}(\gamma)) = \mathcal{E}(\gamma)\}.$$

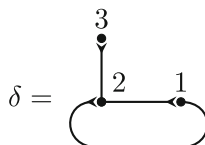
We now define the *labeling* of graph $\gamma \in \Gamma$ as a bijection $\lambda : [|\gamma|] \rightarrow \mathcal{V}(\gamma)$. By convention, we number the roots first. In particular, *for trees, the root will have number one*. Incidentally, a similar labeling is chosen in the proof of [24, Theorem 7.3].

In the rest of this section we choose one fixed labelling for all aromatic forests. Identifying $\mathcal{V}(\gamma) \equiv [|\gamma|]$ using this labeling, we define, for $\delta \in \Gamma_n^r$, the polynomial vector field $f_\delta \in X^n$ by

$$f_\delta^j(x) := \frac{1}{(\sigma(\delta))^{1/|\delta|}} \prod_{i \in \mathcal{P}(j)} x_i, \tag{13}$$

where an empty product is defined as 1. By construction, $\delta = \text{dep}(f_\delta)$.

As an example, consider the following labeled aromatic molecule



Then

$$f_\delta(x) = \begin{pmatrix} x_2 \\ x_1 x_3 \\ 1 \end{pmatrix}.$$

Lemma 4.7 *If $\delta = \delta_1 \delta_2$ then*

$$f_\delta = f_{\delta_1} \oplus f_{\delta_2}.$$

Proof Clear from the definition of f_δ . □

Lemma 4.8 *Let $\mu, \mu' \in M \subset \Gamma^0$ be aromatic molecules and $\tau, \tau' \in T \subset \Gamma^1$ trees. The elementary differentials of μ and τ applied to $f_{\mu'}$ and $f_{\tau'}$ are given by*

$$\begin{aligned} \mu[f_{\tau'}] &= 0 \\ \mu[f_{\mu'}] &= \begin{cases} 1 & \text{if } \mu = \mu' \\ 0 & \text{otherwise} \end{cases} \\ \tau[f_{\mu'}] &= 0 \\ \tau[f_{\tau'}]^1 &= \begin{cases} 1 & \text{if } \tau = \tau' \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Proof Consider two connected graphs $\gamma, \gamma' \in \Gamma$. Define by f the product in (13) for the graph γ' , i.e, $f := (\sigma(\delta))^{1/|\delta|} f_{\gamma'}$. Let \mathcal{P} and \mathcal{P}' denote the parent functions in the graphs γ and γ' . From (9) we find

$$\begin{aligned} \gamma[f] &= \sum_{v \in \text{hom}(\gamma, \gamma')} \prod_{v \in \mathcal{V}(\gamma)} \frac{\partial^{|\mathcal{P}(v)|}}{\partial x_{v(\mathcal{P}(v))}} (f)^{v(v)} \prod_{r \in \mathcal{R}(\gamma)} \partial_{v(r)} \\ &= \sum_v \prod_{v \in \mathcal{V}(\gamma)} \frac{\partial^{|\mathcal{P}(v)|}}{\partial x_{v(\mathcal{P}(v))}} \prod_{i \in \mathcal{P}'(v(v))} x_i \prod_{r \in \mathcal{R}(\gamma)} \partial_{v(r)} \\ &= \sum_v \prod_{v \in \mathcal{V}(\gamma)} \prod_{i \in \mathcal{P}'(v(v)) \setminus v(\mathcal{P}(v))} x_i \prod_{r \in \mathcal{R}(\gamma)} \partial_{v(r)}. \end{aligned}$$

If $v \in \text{hom}(\gamma, \gamma')$ sends more than one edge in γ to the same edge in γ' , the expression becomes 0, so it is enough to consider graph embeddings, $v \in (\gamma \hookrightarrow \gamma')$, the maps that are injective both on vertices and edges. If some edge in γ' is not covered by the image of an edge in γ , the result is a monomial $\prod_i x_i$ running over all edges not covered by the embedding, which evaluates to 0 at $x = 0$. If v is a graph isomorphism the expression evaluates to 1. Hence, we conclude that for the root component (numbered one by convention):

$$(\gamma[f])^1(0) = \begin{cases} \sigma(\gamma) & \text{if } \gamma = \gamma', \\ 0 & \text{otherwise.} \end{cases}$$

□

4.3 Aromatic series on dual vector fields

For regular B-series, Lemma 4.8 provides a dual basis to the elementary differential. For aromatic series we must take into account polynomial relations, such as $(\mu\mu)[f_\mu] = (\mu[f_\mu])^2$ for $\mu \in M$. The goal of this section is thus to construct the equivalent of the vector fields of Lemma 4.8, but for aromatic trees. As we shall see, we cannot achieve a corresponding result, but a result that suffices for our purpose.

We need an elementary result first. If $\gamma \in \Gamma$ is disconnected, $\gamma = \gamma_1 \gamma_2$, we can decompose $\gamma[f]$ in the following way.

Lemma 4.9 *For $\gamma_1 \in \Gamma^{r_1}$, $\gamma_2 \in \Gamma^{r_2}$ we have*

$$\gamma_1 \gamma_2[f] = \gamma_1[f] \cdot \gamma_2[f]$$

where the product on the right denotes the symmetric tensor product.

Proof If the graph $\gamma = \gamma_1 \gamma_2$ is disconnected then $\mathcal{P}(v_1) \subset \mathcal{V}(\gamma_1)$ for all $v_1 \in \mathcal{V}(\gamma_1)$ and similar for γ_2 , and the result is readily checked from (9). \square

Note that we will only use that result for products of graphs in Γ^0 , i.e., products of molecules, or products of elements in Γ^0 and Γ^1 . In particular, the tensor product on the right will always be either a scalar or a vector.

We now come to the central result of this section.

Lemma 4.10 *Fix aromatic molecules $\mu_1, \dots, \mu_m \in M$, scalars $\lambda_1, \dots, \lambda_m \in \mathbf{R}$, and a tree $\tau \in T$. Define*

$$f := f_\tau \oplus \lambda_1 f_{\mu_1} \oplus \dots \oplus \lambda_m f_{\mu_m}.$$

Let π be the projection on the first components, that is, $\pi f = f_\tau$. Choose an arbitrary element $\gamma \in \Gamma^1$. If $\gamma = \mu_1^{p_1} \dots \mu_m^{p_m} \tau$ for some integers $p_1, \dots, p_m \geq 0$, then

$$\pi(\gamma[f]) = \lambda_1^{|\mu_1|p_1} \dots \lambda_m^{|\mu_m|p_m} \partial_1,$$

otherwise, $\pi(\gamma[f]) = 0$.

Proof An aromatic tree γ can always be written as

$$\gamma = \sigma \mu_1^{p_1} \dots \mu_m^{p_m} \tau'$$

for some integers $p_i \geq 0$, an element $\sigma \in \Gamma^0$ which does not contain any of the molecules μ_i , and a regular tree $\tau' \in T$.

First, using Lemma 4.9 we obtain that

$$\gamma[f] = \sigma[f] (\mu_1[f])^{p_1} \dots (\mu_m[f])^{p_m} \tau'[f] \tag{14}$$

Now, using Lemma 4.5 and that μ is $|\mu|$ -linear, we obtain that for any molecule $\mu \in M$:

$$\mu[f] = \mu[f_\tau] + \lambda_1^{|\mu|} \mu[f_{\mu_1}] + \dots + \lambda_m^{|\mu|} \mu[f_{\mu_m}]. \tag{15}$$

Using (15) with $\mu = \mu_i$, we obtain using Lemma 4.8 that $\mu_i[f] = \lambda_i^{|\mu_i|}$. If σ is not empty, it contains one molecule μ , which, by assumption is distinct from any of the μ_i , and (15) is then zero. As (15) factors $\sigma[f]$ which in turn factorises (14), the whole expression (14) is zero. If σ is empty, we have by convention that $\sigma[f] = 1$. Similarly, we obtain from Lemma 4.5 that

$$\tau'[f] = \tau'[f_\tau] \oplus \lambda_1^{|\tau'|} \tau'[f_{\mu_1}] \oplus \dots \oplus \lambda_m^{|\tau'|} \tau'[f_{\mu_m}].$$

We conclude that if $\tau' \neq \tau$ then the root component (which, by convention, has label 1) of (14) is zero. If $\tau' = \tau$ then $\tau'[f]^1 = \tau[f]^1 = 1$, which concludes the proof. \square

5 Proof of the core result

We now set out to prove what is the core result of this paper. Indeed, the following result is the main ingredient in the proof of Theorem 2.4.

Theorem 5.1 *For any degree k , \mathcal{F} induces a bijection between elements of $\langle T_k \rangle$ and equivariant sequences of k -forms.*

Theorem 5.1 is proved through Proposition 5.2 (injectivity), Proposition 5.3 (compatibility), and Proposition 5.6 (surjectivity).

5.1 Injectivity

Proposition 5.2 *The elementary differential map \mathcal{F} is injective.*

Proof Suppose that $\gamma \in \langle \Gamma_k^1 \rangle$ and that $\mathcal{F}(\gamma) = 0$. Then we have in particular $\mathcal{F}_k(\gamma) = 0$ and Proposition 3.2 yields $\gamma = 0$. \square

5.2 B-series are equivariant sequences

In this subsection we prove the following result.

Proposition 5.3 *\mathcal{F} maps $\langle T_k \rangle$ to the space of equivariant sequences of k -forms.*

The geometry of affine spaces is closely related to the existence of a flat, torsion-free connection. Note that for each $n \in \mathbb{N}$ the following connection $\triangleright_n : X^n \times X^n \rightarrow X^n$ is well defined because its result is also a polynomial vector field:

$$(f \triangleright_n g)(x) := \left. \frac{\partial}{\partial t} \right|_{t=0} g(f(x + tf(x))).$$

As before, we consider now a connection \triangleright as a *sequence* of the connections on all dimensions n . We also omit the dimension when the context is clear, so we write:

$$f \triangleright g := f \triangleright_n g \quad f, g \in X^n.$$

Lemma 5.4 *The connection \triangleright is an equivariant sequence in the following sense:*

$$f \overset{A}{\rightsquigarrow} \tilde{f} \text{ and } g \overset{A}{\rightsquigarrow} \tilde{g} \implies (f \triangleright g) \overset{A}{\rightsquigarrow} (\tilde{f} \triangleright \tilde{g}) \quad A \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m).$$

Proof Let $y = Ax$. Then

$$\begin{aligned} A(f \triangleright g)(x) &= \left. \frac{\partial}{\partial t} \right|_{t=0} Ag(x + tf(x)) = \left. \frac{\partial}{\partial t} \right|_{t=0} \tilde{g}(Ax + tAf(x)) \\ &= \left. \frac{\partial}{\partial t} \right|_{t=0} \tilde{g}(y + t\tilde{f}(y)) = (\tilde{f} \triangleright \tilde{g})(y). \end{aligned}$$

□

In the language of algebra, $\{X^n, \triangleright_n\}$ is an example of a *pre-Lie algebra* [7], i.e., a vector space with a bilinear binary operation $\triangleright = \triangleright_n$ that is neither commutative nor associative, but satisfies the pre-Lie relation

$$f \triangleright (g \triangleright h) - (f \triangleright g) \triangleright h = g \triangleright (f \triangleright h) - (g \triangleright f) \triangleright h.$$

Recall that T denotes the set of all rooted trees, that for $\tau \in T$, $|\tau|$ denote the number of vertices in τ and that $\langle T \rangle$ is the free \mathbf{R} -vector space over T .

The *free pre-Lie algebra*, denoted $\{\langle T \rangle, \triangleright\}$, is defined by the binary operation $\triangleright: \langle T \rangle \times \langle T \rangle \rightarrow \langle T \rangle$ given by *grafting* on trees. That is, the binary operation \triangleright given by summing over all trees resulting from attaching successively the tree τ_1 to each vertex of τ_2 :

$$\tau_1 \triangleright \tau_2 := \sum_{v \in V(\tau_2)} \tau_1 \circ_v \tau_2,$$

where $\tau_1 \circ_v \tau_2$ denotes attachment of the root of τ_1 to the vertex v of τ_2 via a new edge. The free pre-Lie algebra is *universal* in the category of pre-Lie algebras:

Proposition 5.5 [9] *For any pre-Lie algebra $\{\mathcal{A}, \triangleright\}$ and any $f \in \mathcal{A}$, there exists a unique map $\mathcal{F}(\cdot)[f]: \langle T \rangle \rightarrow \mathcal{A}$ defined by linearity and the recursion*

$$\begin{aligned} \mathcal{F}(\bullet)[f] &= f \\ \mathcal{F}(\tau_1 \triangleright \tau_2)[f] &= \mathcal{F}(\tau_1)[f] \triangleright \mathcal{F}(\tau_2)[f], \quad \tau_1, \tau_2 \in \langle T \rangle. \end{aligned} \tag{16}$$

When $\mathcal{A} = X^n$ and $\tau \in T$, the elements $\mathcal{F}(\tau)[f] \in X^n$ are thus the elementary differentials that we defined in Sect. 3.6, as the recursion equations (3) are fulfilled in that case.

The following result establishes a proof of Proposition 5.3.

Proof of Proposition 5.3 By Proposition 5.5, we can express the elementary differential map \mathcal{F} using connections and the tree \bullet . As $\mathcal{F}(\bullet)[\cdot]$ is obviously an equivariant sequence, and as the connection is equivariant in the sense of Lemma 5.4, we conclude that $\mathcal{F}(\tau)[\cdot]$ also is, for any tree $\tau \in T$. □

5.3 Surjectivity

The main result of this subsection is the following result.

Proposition 5.6 \mathcal{F} is a surjection from $\langle T_k \rangle$ to the space of equivariant sequences of k -forms.

Proof The proof contains two steps: assuming P is an equivariant sequence of k -forms, we prove that there exists a $\gamma \in \langle \Gamma_k^1 \rangle$ such that $P = \mathcal{F}(\gamma)$ (Lemma 5.7), then we prove that γ must in fact be in $\langle T_k \rangle$ (Lemma 5.8). □

Lemma 5.7 \mathcal{F} is a surjection from $\langle \Gamma_k^1 \rangle$ to the space of equivariant sequences of k -forms.

Proof Let P be an equivariant sequence of k -forms. For any n , using Proposition 3.2 we have $\gamma_n \in \langle \Gamma_k^1 \rangle$ such that $\mathcal{F}_n(\gamma_n) = P_n$. We aim to show that $P = \mathcal{F}(\gamma_k)$.

1. We first show that for $m \leq n$, we have

$$\mathcal{F}_m(\gamma_m) = \mathcal{F}_m(\gamma_n). \tag{17}$$

Indeed, take a $g \in X^m$ and construct $f = g \oplus 0$. The equivariance property Lemma 4.2 implies that $P_n[g \oplus 0] = P_m[g] \oplus 0$, which gives $\mathcal{F}_n(\gamma_n)[g \oplus 0] = \mathcal{F}_m(\gamma_m)[g] \oplus 0$. Using Lemma 4.6 we obtain $\mathcal{F}_m(\gamma_n)[g] \oplus 0 = \mathcal{F}_m(\gamma_m)[g] \oplus 0$, so (1) is proved.

2. In particular, for $n \leq k$, (1) gives $\mathcal{F}_n(\gamma_k) = \mathcal{F}_n(\gamma_n) = P_n$. For $k \leq n$ (1) gives $\mathcal{F}_k(\gamma_k) = \mathcal{F}_k(\gamma_n)$, but as γ_n is a k -form, we can use Proposition 3.2 and obtain $\gamma_k = \gamma_n$. For $n \geq k$, this gives $P_n = \mathcal{F}_n(\gamma_k)$.

We have shown that $\mathcal{F}_n(\gamma_k) = P_n$ for any n , so we conclude that $P = \mathcal{F}(\gamma_k)$. □

Lemma 5.8 Let P be an equivariant sequence of k -forms. If $\mathcal{F}(\gamma) = P$ for $\gamma \in \langle \Gamma_k^1 \rangle$, then $\gamma \in \langle T_k \rangle$.

Proof Fix a tree $\tau \in T$. The element $\gamma \in \langle \Gamma_k^1 \rangle$ can be written as

$$\gamma = p(\mu_1, \dots, \mu_m)\tau + \gamma'$$

where γ' does not contain any occurrence of the tree τ , and p is a polynomial over some molecules μ_1, \dots, μ_n . For instance, if $\gamma = (3(\mu_1)^2\mu_2 + (\mu_1)^3)\tau + \gamma'$, then $p(X_1, X_2) = 3X_1^2X_2 + X_1^3$. The goal is now to prove that p is constant. Recall the notations of Lemma 4.10, in particular the special vector $f := f_\tau \oplus \lambda_1 f_{\mu_1} \oplus$

$\dots \oplus \lambda_m f^{\mu_m}$ and the projection π which projects on the first components of the decomposition of f . Using Lemma 4.10 we have

$$\pi \gamma[f] = p(\lambda_1^{|\mu_1|}, \dots, \lambda_m^{|\mu_m|}) \partial_1. \tag{18}$$

Now, using that the sequence η is equivariant, that is, using Lemma 4.2 for $f = f_\tau \oplus g$, one obtains $\pi \gamma[f] = \gamma[f_\tau]$. Now, using Lemma 4.10, we obtain $\gamma[f_\tau] = p(0, \dots, 0) \partial_1$, so

$$\pi \gamma[f] = p(0, \dots, 0) \partial_1. \tag{19}$$

We deduce from (19) and (18) that as $p(\lambda_1^{|\mu_1|}, \dots, \lambda_m^{|\mu_m|}) = p(0, \dots, 0)$, the polynomial p must be constant. Finally, we conclude that γ is in $\langle T_k \rangle$. \square

6 Transfer argument

6.1 Transfer to the Taylor terms

When dealing with B-series, one treats each order separately, so that statements are stable with respect to truncations (or, put differently, with respect to the inverse limit topology). The terms in a B-series correspond to the terms in the Taylor expansion of a map, so statements about B-series imply statements about Taylor terms and vice versa. Notice, however, that Theorem 2.4 says something about the family ϕ as a whole, not about its Taylor terms. Therefore, to prove Theorem 2.4 through Theorem 5.1, we need to show that a Taylor term of an affine equivariant integrator map is also an affine equivariant integrator map; we call this a *transfer argument*.

We first show that affine equivariant sequences preserve fixed points:

Lemma 6.1 *If ϕ is an affine equivariant sequence of smooth maps, then for any integer n , any point $x \in \mathbf{R}^n$, and any vector field $f \in \mathfrak{X}_0(\mathbf{R}^n)$, we have*

$$f(x) = 0 \implies \phi_n(f)(x) = 0.$$

This implies in particular $\phi_n(0) = 0$.

Proof Consider the trivial affine injection $\mathbf{R}^0 \rightarrow \mathbf{R}^n$ defined by $a(0) = x$. The mapping a intertwines the zero vector field on \mathbf{R}^0 and f , so by equivariance of ϕ_n , we have: $\phi_n(f)(x) = 0$.

Let $\phi_n: \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$ be a smooth map, and consider, as in Sect. 2, the k th Taylor coefficient $D^k \phi_n(0)$. Recall that this is a symmetric multi-linear, vector valued map. The corresponding homogeneous polynomial $\Delta_k(\phi_n): \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$ is

$$\Delta_k(\phi_n)(f) := D^k \phi_n(0)[f, \dots, f].$$

The following transfer argument is the main result of this section.

Proposition 6.2 *Let $\phi = \{\phi_n\}_{n \in \mathbf{N}}$ be an affine equivariant sequence of smooth maps. Then, for any fixed $k \in \mathbf{N}$, the family of maps $\{\Delta_k(\phi_n)\}_{n \in \mathbf{N}}$ is an affine equivariant sequence of homogeneous polynomials of degree k .*

Proof First, by Lemma 6.1, we have that $\phi_n(0) = 0$. The Taylor polynomial $\Delta_k(\phi_n)$ for a smooth map $\phi_n: \mathfrak{X}_0(\mathbf{R}^n) \rightarrow \mathfrak{X}_0(\mathbf{R}^n)$ such that $\phi_n(0) = 0$ is given by [20, §5.11]

$$\Delta_k(\phi_n)(f) = \partial_{t_1} \dots \partial_{t_k} \phi_n((t_1 + \dots + t_k)f) \Big|_{t_1=0, \dots, t_k=0} \tag{20}$$

Consider two vector fields $f \in \mathfrak{X}_0(\mathbf{R}^n)$ and $g \in \mathfrak{X}_0(\mathbf{R}^m)$, related by an affine map $a(x) = Ax + b$, i.e., $f \overset{a}{\rightsquigarrow} g$. Then $(\lambda f) \overset{a}{\rightsquigarrow} (\lambda g)$ for any $\lambda \in \mathbf{R}$. Since ϕ is an equivariant sequence, we have $\phi_n(\lambda f) \overset{a}{\rightsquigarrow} \phi_m(\lambda g)$, which, by its definition (2), means $\phi_m(\lambda g)(ax) = A\phi_n(\lambda f)(x)$ for $x \in \mathbf{R}^n$. Therefore, taking $\lambda = t_1 + \dots + t_n$, we obtain

$$\phi_m((t_1 + \dots + t_k)g)(ax) = A\phi_n((t_1 + \dots + t_k)f)(x).$$

We conclude, using the defining property of Δ_k in (20), that $\Delta_k(\phi_n)(f) \overset{a}{\rightsquigarrow} \Delta_k(\phi_m)(g)$. □

6.2 Extension principle

We now observe that by Lemma 6.1, and because we have polynomials instead of nonlinear maps, we can apply Peetre’s theorem and assert that ϕ_n depends on a finite number of derivatives of the vector field. Furthermore, by equivariance at each dimension, we can restrict the study to that of (linear) equivariant sequences of k -forms on polynomial vector fields as defined in Sect. 3.4.

We only give a sketch of the proof, as the details are similar to [24], the main novelty being using sequences instead of maps.

Proposition 6.3 *There is a bijection between the space of affine equivariant sequences of homogeneous polynomials of degree k on vector fields, and the space of (linear) equivariant sequences of k -forms on polynomial vector fields.*

Proof Observe that Lemma 6.1 implies in particular that ϕ_n is support non-increasing, or *local* (see [24, §2.3]) so we can use the extension principle in [24, Proposition 4.2] We thus obtain for each dimension n a map $\bar{\varphi}_n$ defined from polynomial vector fields to \mathbf{R}^n ; moreover, this map is $\mathbf{GL}(n)$ -equivariant. The relation between φ and $\bar{\varphi}$ is $\varphi(f)(x) = \bar{\varphi}(T(f))(0)$, where T is the Taylor development of f . Note that $\bar{\varphi}$ has finite order, i.e., it only needs the Taylor development up to some finite order k .

Consider an affine equivariant sequence φ_n of local k -forms on vector fields. We thus obtain a sequence $\bar{\varphi}_n$ of $\mathbf{GL}(n)$ equivariant maps. As the sequence φ_n is affine equivariant, we obtain that the sequence $\bar{\varphi}_n$ is equivariant in the sense of Sect. 3.4.

On the other hand, given an equivariant sequence in the sense of Sect. 3.4, we obtain a sequence of affine-equivariant maps φ_n . It is then straightforward to check that this sequence is in fact an equivariant sequence. □

Acknowledgments Klas Modin is supported by the Swedish Research Council (Contract VR-2012-335). Olivier Verdier is supported by the J.C. Kempe memorial fund (Grant No. SMK-1238). Robert McLachlan is supported by the Marsden Fund of the Royal Society of New Zealand. We express our greatest gratitude to the referees for their numerous useful remarks, in particular for noticing that one can avoid the assumption of locality. We are also grateful to Alexander Schmeding for giving us the right functional analytic framework of compactly supported vector fields in the definition of an integrator.

References

1. Bochev, P.B., Clint, S.: On quadratic invariants and symplectic structure. *BIT Numer. Math.* **34**(3), 337–345 (1994)
2. Brouder, C.: Runge–Kutta methods and renormalization. *Eur. Phys. J. C-Part. Fields* **12**(3), 521–534 (2000)
3. Brouder, C.: Trees, renormalization and differential equations. *BIT Numer. Math.* **44**(3), 425–438 (2004)
4. Butcher, J.C.: Coefficients for the study of Runge–Kutta integration processes. *J. Aust. Math. Soc.* **3**(02), 185–201 (1963)
5. Butcher, J.C.: An algebraic theory of integration methods. *Math. Comput.* **26**(117), 79–106 (1972)
6. Butcher, J.C.: *Numerical Methods for Ordinary Differential Equations*. Wiley, New York (2008). (ISBN 9780470753750)
7. Calaque, D., Ebrahimi-Fard, K., Manchon, D.: Two interacting Hopf algebras of trees: a Hopf-algebraic approach to composition and substitution of B-series. *Adv. Appl. Math.* **47**(2), 282–308 (2011)
8. Cayley, A.: On the theory of analytic forms called trees. *Philos. Mag.* **XIII**, 172–176 (1857)
9. Chapoton, F., Livernet, M.: Pre-lie algebras and the rooted trees operad. *Int. Math. Res. Not.* **2001**(8), 395–408 (2001)
10. Chartier, P., Hairer, E., Vilmart, G.: Algebraic structures of B-series. *Found. Comput. Math.* **10**(4), 407–427 (2010)
11. Chartier, P., Murua, A.: An algebraic theory of order. *ESAIM Math. Model. Numer. Anal.* **43**(04), 607–630 (2009)
12. Chartier, P., Faou, E., Murua, A.: An algebraic approach to invariant preserving integrators: the case of quadratic and Hamiltonian invariants. *Numer. Math.* **103**(4), 575–590 (2006)
13. Connes, A., Kreimer, D.: Lessons from quantum field theory: Hopf algebras and spacetime geometries. *Lett. Math. Phys.* **48**(1), 85–96 (1999)
14. Gubinelli, M.: Ramification of rough paths. *J. Differ. Equ.* **248**(4), 693–721 (2010)
15. Hairer, E., Wanner, G.: On the Butcher group and general multi-value methods. *Computing* **13**(1), 1–15 (1974)
16. Hairer, E., Wanner, G.: *Solving Ordinary Differential Equations. II: Stiff and Differential-Algebraic Problems*. Springer, Berlin (2010). doi:10.1007/978-3-642-05221-7. (ISBN 978-3-642-05221-7)
17. Hairer, E., Zbinden, C.J.: On conjugate symplecticity of B-series integrators. *IMA J. Numer. Anal.* **33**(1), 57–79 (2013)
18. Huta, A.: Une amélioration de la méthode de Runge–Kutta–Nyström pour la résolution numérique des équations différentielles du premier ordre. *Acta Fac. Nat. Univ. Comenian. Math.* **1**, 201–224 (1956)
19. Iserles, A., Quispel, G.R.W., Tse, P.: B-series methods cannot be volume-preserving. *BIT Numer. Math.* **47**(2), 351–378 (2007). doi:10.1007/s10543-006-0114-8
20. Kriegl, A., Michor, P.W.: *The Convenient Setting of Global Analysis*. AMS Mathematical Surveys and Monographs. American Mathematical Society, Providence (1997). (ISBN 9780821807804)
21. McLachlan, R.I., Quispel, G.R.: Six lectures on the geometric integration of ODEs. In: *Foundations of Computational Mathematics*. London Mathematical Society Lecture Note Series, pp. 155–210. Cambridge University Press, Cambridge (2001a)
22. McLachlan, R.I., Quispel, G.R.W.: What kinds of dynamics are there? Lie pseudogroups, dynamical systems and geometric integration. *Nonlinearity* **14**(6), 1689–1705 (2001b)
23. Merson, R.H.: *Data processing and automatic computing machines: proceedings of a conference held at Weapons Research Establishment, Salisbury, S.A., June 3rd–8th*, pp. 1101-1-110-22. Weapons Research Establishment, Salisbury, South Australia (1957)
24. Munthe-Kaas, H.Z., Verdier, O.: *Aromatic Butcher Series*. Foundations of Computational Mathematics (2015). doi:10.1007/s10208-015-9245-0. arXiv:1308.5824

25. Murua, A.: The Hopf algebra of rooted trees, free Lie algebras, and Lie series. *Found. Comput. Math.* **6**(4), 387–426 (2006)
26. Quispel, G.R.W., McLaren, D.I.: A new class of energy-preserving numerical integration methods. *J. Phys. A Math. Theor.* **41**(4), 7 (2008). doi:[10.1088/1751-8113/41/4/045206](https://doi.org/10.1088/1751-8113/41/4/045206). (ISSN 1751-8113; 1751-8121/e)