B-series tutorial

Anne Kværnø

February 15, 2008

1 Introduction

This note is written to explain how to develop an order theory for onestep methods applied to initial value problems, in particular ordinary differential equations (ODEs) and differential-algebraic equations (DAEs). In modern advanced text-books, like the ones in the reference lists, there are excellent expositions of the theory for Runge-Kutta methods applied to ODEs, and some even treats DAE problems. In the literature, order theory for a wide range of problems and one-step methods has been treated. But mostly, one combination at the time. The ambition for this note is explain how to develop order theory, from some rather general assumptions. These assumptions are essentially that both the exact and the numerical solution can be written, as least formally, as power series of the stepsize h, and that each term corresponding to a certain power of h might be split into several terms. But exactly how these series looks like depends on both the problem and the method under consideration.

Section 2 presents some general definitions and results needed for the rest of the paper. These results are not related to a certain problem or method. Section 3 will use these results to develop order theory for Runge-Kutta methods applied to ODEs. A later section (not yet finished) will consider Runge-Kutta methods applied to DAEs. The paper might be extended with more sections in some remote future.

2 Key definitions and results

The aim of this paper is to develop an order theory for some initial value problem solved by some one-step method. To do so, we will assume that both the exact solution, $y(t_0+h)$ satisfying the initial condidtion $y(t_0) = y_0$, and the numerical solution $y_1 \approx y(t_0+h)$ can be written as power series of h, thus

$$y(t_0+h) = \sum_{p=0}^{\infty} y^{(p)}(0)\frac{h^p}{p!} \tag{1}$$

$$y_1 = y_1(h) = \sum_{p=0}^{\infty} y_1^{(p)}(0)\frac{h^p}{p!} \tag{2}$$

where $y^{(p)}(0) = \dfrac{\partial^p y(t_0+h)}{\partial h^p}|_{h=0} = \dfrac{\partial^p y(t)}{\partial t^p}|_{t=t_0}$ and $y_1^{(p)}(0) = \dfrac{\partial^p y_1(h)}{\partial h^p}|_{h=0}$. The local truncation error d is given by

$$d = y(t_0+h) - y_1(h) = \sum_{p=0}^{\infty}(y^{(p)}(0) - y_1^{(p)}(0))\frac{h^p}{p!}$$

thus we will find the order of a method by comparing $y^{(p)}(0)$ with $y_1^{(p)}(0)$. However, each of these terms might split into several terms. We will use *trees*, denoted by τ to keep track of these terms. Doing so, $y^{(p)}$ can be written as

$$y^{(p)}(0) = \sum_{\substack{\tau\in T\\|\tau|=p}} \alpha(\tau)F(\tau)(y_0) \tag{3}$$

where *the elementary differentials* are those mysterious terms we are looking for, while $\alpha(\tau)$ are the number of equal terms in the sum. The meaning of *the order of the tree*, $|\tau|$, should be clear. Using this notation, the exact solution is written as the following power series of h:

$$y(t_0+h) = \sum_{\tau\in T}\alpha(\tau)F(\tau)(y_0)\frac{h^{|\tau|}}{|\tau|!} \tag{4}$$

The standard convention is to use $\emptyset$ for the order zero tree, so for $p = 0$ (3) becomes

$$y(t_0) = \sum_{\substack{\tau\in T\\|\tau|=0}} \alpha(\tau)F(\tau)(y_0) = \alpha(\emptyset)F(\emptyset)(y_0).$$

We can immediately conclude that $|\emptyset| = 0$, $\alpha(\emptyset) = 1$ and $F(\emptyset)(y_0) = y_0$. The remaining terms will be found by recursion. We will like the numerical solution y_1 to be written in terms of a similar series, that is

$$y_1(h) = \sum_{\tau\in T}\alpha(\tau)\psi(\tau)F(\tau)(y_0)\frac{h^{|\tau|}}{|\tau|!} \tag{5}$$

where the coefficients $\psi(\tau)$ are method dependent. This motivates the following definition:

Definition 1 *A B-series is a series of the form*

$$B(\phi, y_0; h) = \sum_{\tau\in T}\alpha(\tau)\phi(\tau)F(\tau)(y_0)\frac{h^{|\tau|}}{|\tau|!}$$

where $\alpha(\tau) : T \to \mathbb{N}\backslash\{0\}$, $|\tau| : T \to \mathbb{N}$, $\phi(\tau) : T \to \mathbb{R}$.

We will call the series *consistent* if $B(\phi, y_0; 0) = y_0$, or equivalent

$$B(\phi, y_0; h) = y_0 + \sum_{\substack{\tau\in T\\|\tau|\geq 1}} \alpha(\tau)\phi(\tau)F(\tau)(y_0)\frac{h^{|\tau|}}{|\tau|!}.$$

Remark 1: For the exact solution of the problem under consideration the B-series will be specified by $y(t_0+h) = B(1, y_0; h)$ and we hope to find $\psi(\tau)$ such that $y_1 = B(\psi, y_0; h)$.
Remark 2: Definition 1 is less specific than what is common in the literature, where B-series is strictly related to ODEs, such that T, α and the elementary differentials F are defined. But our long-term objective is to develop order theory for a lot of different problem and methods, and for this purpose an open definition is more convenient.

The next definition can be found in [3], see also Butcher [1]:

Definition 2 *Let* $y = [y_1, y_2, \cdots, y_N]^T \in \mathbb{R}^N$ *and* $f(y) = [f_1(y), f_2(y), \cdots, f_n(y)]^T \in \mathbb{R}^n$*. The* m'*th Frechet derivative of* f*, denoted by* $f_{my}(y)$ *is a* m*-linear operator* $\mathbb{R}^N \times \mathbb{R}^N \times \cdots \times \mathbb{R}^N$ *(*m *times)* $\to \mathbb{R}^n$*. Evaluation of component* i *of this operator working on the* m *operands* $v_1, v_2, \cdots v_m \in \mathbb{R}^N$ *is given by*

$$[f_{my}(y)(v_1, v_2, \cdots, v_m)]_i = \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} \cdots \sum_{j_m=1}^{N} \frac{\partial^m f_i(y)}{\partial y_{j_1} \partial y_{j_2} \cdots \partial y_{j_m}} v_{1,j_1} v_{2,j_2} \cdots v_{m,j_m}$$

where $v_l = [v_{l,1}, v_{l,2}, \cdots v_{l,N}] \in \mathbb{R}^N$ *for* $l = 1, 2, \cdots, m$*.*

Note that the m'th Frechet derivative is independent of permutations of its operands, thus e.g. $f_{3y}(v_1, v_2, v_3) = f_{3y}(v_3, v_1, v_2)$.

In the next lemma and its proof the following concepts are useful:

A *list* of trees, denoted by $\{\tau_1, \tau_2, \cdots, \tau_m\}$, $\tau_i \in T$, $i = 1, \cdots, m$ is a ordered set of trees, where each tree might appear more than once. If $\tau_1, \tau_2 \in T$ then $\{\tau_1, \tau_2, \tau_1\}$ and $\{\tau_2, \tau_1, \tau_1\}$ are two different lists. If a tree appear k times in the list, the tree has multiplicity k. A *multiset* of trees, denoted by $(\tau_1, \tau_2, \cdots, \tau_m)$ is a set of trees where multiplicity is allowed but order does not matter. So $(\tau_1, \tau_2, \tau_1) = (\tau_2, \tau_1, \tau_1)$. A tree with multiplicity k will sometimes be denoted by t^k, so $(\tau_1, \tau_2, \tau_1) = (\tau_1^2, \tau_2)$. The set of all possible lists of trees is denoted $\tilde{U}$, and the set of all possible multisets is denoted U:

$$\begin{aligned}
\tilde{U} &= \{\{\tau_1, \tau_2, \cdots, \tau_m\} : \tau_i \in T, \quad i = 1, \cdots, m, \quad m = 0, 1, 2, \cdots\}, \\
U &= \{(\tau_1, \tau_2, \cdots, \tau_m) : \tau_i \in T, \quad i = 1, \cdots, m, \quad m = 0, 1, 2, \cdots\}.
\end{aligned}$$

The following lemma is the key tool in this paper:

Lemma 1 *Let* $Z = B(\phi, y_0; h)$ *be a consistent B-series and let* $f : \mathbb{R}^N \to \mathbb{R}^n$*. Then*

$$f(Z) = \sum_{u \in U} \alpha_f(u) \beta_f(u) F_f(u)(y_0) \frac{h^{|u|_f}}{|u|_f!}$$

where for each $u = (\tau_1, \tau_2, \cdots, \tau_m) \in U$ *we have*

$$|\emptyset|_f = 0, \qquad |u|_f = \sum_{i=1}^{m} |\tau_i| \qquad \big(\textit{ and } |u| \geq 1 \textit{ if } u \neq (\emptyset)\big).$$

$$\alpha_f((\emptyset)) = 1, \qquad \alpha_f(u) = |u|_f! \prod_{i=1}^{m} \frac{\alpha(\tau_i)}{|\tau_i|!} \prod_{j=1}^{l} \frac{1}{k_j!}.$$

$$\beta_f((\emptyset)) = 1, \qquad \beta_f(u) = \prod_{i=1}^{m} \phi(\tau_i)$$

$$F_f((\emptyset))(y_0) = f(y_0), \quad F_f(u)(y_0) = f_{my}(y_0)(F(\tau_1)(y_0), F(\tau_2)(y_0), \cdots, F(\tau_m)(y_0)).$$

In the expression for $\alpha_f(u)$*,* u *is a multiset with* l *distinct trees, each of multiplicity* k_j*,* $j = 1, \cdots, l$*.*

Proof: Taylors theorem for a multivalued function $f(y)$, $f : \mathbb{R}^N \to \mathbb{R}^n$ states that if $\delta \in \mathbb{R}^N$ then

$$f(y + \delta) = f(y) + f_y(y)\delta + \frac{1}{2} f_{2y}(\delta, \delta) + \cdots = \sum_{m=0}^{\infty} \frac{1}{m!} f_{my}(y)(\overbrace{\delta, \delta, \cdots, \delta}^{m \text{ times}}).$$

Since $\phi(\emptyset) = 1$ we have

$$\begin{aligned} f(Z) &= f\Big(y_0 + \sum_{\substack{\tau \in T \\ |\tau| \geq 1}} \alpha(\tau)\phi(\tau)F(\tau)(y_0)\Big) \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} f_{my}(y_0)\Big(\sum_{\substack{\tau_1 \in T \\ |\tau_1| \geq 1}} \alpha(\tau_1)\phi(\tau_1)F(\tau_1)(y_0)\frac{h^{|\tau_1|}}{|\tau_1|!}, \cdots, \sum_{\substack{\tau_m \in T \\ |\tau_m| \geq 1}} \alpha(\tau_m)\phi(\tau_m)F(\tau_m)(y_0)\frac{h^{|\tau_m|}}{|\tau_m|!}\Big). \end{aligned}$$

Taking advantage of the fact that Frechet derivative is linear in its operand, the sums as well as all the coefficients can be taken outside the derivative:

$$\begin{aligned} f(Z) = \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{\substack{\tau_1 \in T \\ |\tau_1| \geq 1}} \cdots \sum_{\substack{\tau_m \in T \\ |\tau_m| \geq 1}} & \alpha(\tau_1) \cdots \alpha(\tau_m)\phi(\tau_1) \cdots \phi(\tau_m) \frac{h^{|\tau_1|}}{|\tau_1|!} \cdots \frac{h^{|\tau_m|}}{|\tau_m|!} \\ & \cdot f_{my}(y_0)\Big(F(\tau_1)(y_0), \cdots, F(\tau_m)(y_0)\Big) \end{aligned}$$

These sums can be replaced by the sum over all possible lists of subtrees $\tilde{u} = \{\tau_1, \cdots, \tau_m\} \in \tilde{U}$, including the empty list $\{\emptyset\}$ for $m = 0$. But the Frechet derivative is symmetric in its operands, thus all permutations of the trees in a list results in the same expression. If there are l distinct trees in the list, each with multiplicity k_j, $j = 1, \cdots, l$, then there are

$$\frac{m!}{k_1! k_2! \cdots k_l!}$$

possible permutations of the trees. Thus we can take the sum over all multisets $u = (\tau_1, \cdots, \tau_m) \in U$, just keeping in mind that each such set represents

$\frac{m!}{k_1!k_2!\cdots k_l!}$ equal terms. This will finally give us the expression

$$f(Z) = \sum_{u\in U} \Big(\prod_{i=1}^{m} \frac{\alpha(\tau_i)}{|\tau_i|!} \cdot \prod_{j=1}^{l} \frac{1}{k_j!}\Big) \cdot \Big(\prod_{i=1}^{m} \phi(\tau_i)\Big) \cdot f_{my}(y_0)\big(F(\tau_1)(y_0), \cdots, F(\tau_m)(y_0)\big) \cdot h^{(\sum_{i=1}^{m} |\tau_i|)}$$

which proves the theorem.

□

3 Ordinary differential equations and Runge-Kutta methods

Consider the autonomous initial value problem (IVP) for ordinary differential equations (ODEs)

$$y'(t) = f(y(t)), \qquad y(t_0) = y_0, \qquad f, y \in \mathbb{R}^n \tag{6}$$

where f is assumed sufficiently smooth. In this section we will see how to derive a B-series for the exact and numerical solution of this ODE. The books [1, 2, 3] have all excellent expositions of this topic.

Development of the B-series for the exact solution

The B-series for the exact solution is the series for which $\phi(\tau) = 1$, $\forall \tau \in T$, thus

$$y(t_0 + h) = y_0 + \sum_{\substack{\tau\in T\\ |\tau|\geq 1}} \alpha(\tau) F(\tau)(y_0) \frac{h^{|\tau|}}{|\tau|!}. \tag{7}$$

However, at the moment we know nothing about how this series look like, except that the first two terms are y_0 and $hf(y_0)$. So, we can denote the first two terms by

$$\begin{array}{llll} \emptyset \in T, & |\emptyset| = 0, & \alpha(\emptyset) = 1, & F(\emptyset)(y_0) = y_0, \\ \bullet \in T, & |\bullet| = 1, & \alpha(\bullet) = 1, & F(\bullet)(y_0) = f(y_0). \end{array}$$

Insert this into the ODE (6) and we get.

$$y'(t_0 + h) = f(y(t_0 + h)) = f(B(1, y_0; h))$$

Lemma 1 apply to the right hand side. Taking the Taylor-expansion of the left hand side gives

$$\sum_{p=0}^{\infty} y^{(p+1)}(0) \frac{h^p}{p!} = \sum_{u\in U} \alpha_f(u) F_f(u)(y_0) \frac{h^{|u|_f}}{|u|_f!}$$

using $\beta_f(u) = 1$ since $\phi(\tau_i) = 1$ by assumption. Comparing equal powers of h, using (7) gives

$$y^{(p+1)}(0) = \sum_{\substack{\tau\in T\\|\tau|=p+1}} \alpha(\tau)F(\tau)(y_0) = \sum_{\substack{u\in U\\|u|=p}} \alpha_f(u)F_f(u)(y_0)$$

The two last expressions should be identical, term by term. For each tree $\tau \in T$ of order $p+1$ there is a corresponding multiset $u \in U$ of order p. For $u = (\tau_1, \tau_2, \cdots, \tau_m)$ the corresponding tree will be written as $\tau = [\tau_1, \tau_2, \cdots, \tau_m]$. The graphical interpretation of this is that the roots of the subtrees are attached to a new common root. The order of $\tau = [\tau_1, \tau_2, \cdots, \tau_m]$ is $|u|_f + 1$, and $\alpha(\tau) = \alpha_f(u)$, $F(\tau)(y_0) = F_f(u)(y_0)$ are all given by Lemma 1. Thus we can state the following theorem:

Theorem 1 *The exact solution of the ODE $y'(t) = f(y(t))$, $y(t_0) = y_0$ can be expressed as a B-series*

$$y(t_0 + h) = B(1, y_0; h) = \sum_{\tau\in T} \alpha(\tau)F(\tau)(y_0)\frac{h^{|\tau|}}{|\tau|!}$$

where

- $\emptyset \in T$, $\bullet \in T$ *and if* $\tau_1, \tau_2, \cdots, \tau_m \in T$ *then* $\tau = [\tau_1, \tau_2, \cdots, \tau_m] \in T$.
- $|\emptyset| = 0$, $|\bullet| = 1$ *and if* $\tau = [\tau_1, \tau_2, \cdots, \tau_m] \in T$ *then* $|\tau| = \sum_{i=1}^{m} |\tau_i| + 1$.
- $\alpha(\emptyset) = 1$, $\alpha(\bullet) = 1$ *and if* $\tau = [\tau_1, \tau_2, \cdots, \tau_m] \in T$ *is a tree with* l *distinct subtrees each of multiplicity* k_j, $j = 1, \cdots, l$, *then*

$$\alpha(\tau) = (|\tau| - 1)! \prod_{i=1}^{m} \frac{\alpha(\tau_i)}{|\tau_i|!} \prod_{j=1}^{l} \frac{1}{k_j!}.$$

- $F(\emptyset)(y_0) = y_0$, $F(\bullet) = f(y_0)$ *and if*
  $\tau = [\tau_1, \tau_2, \cdots, \tau_m] \in T$ *then*

$$F(\tau)(y_0) = f_{my}(y_0)(F(\tau_1)(y_0), F(\tau_2)(y_0), \cdots, F(\tau_m)(y_0)).$$

This theorem makes it possible to define B-series related to ODEs as series $B(\phi, y_0; h)$, for which T, $|cdot|$, α and F are all given, but with $\phi : T \to \mathbb{R}$. In the remaining part of this section, this is the B-series we talk about!

From Lemma 1 we get for a consistent B-series, (and for f from the ODE):

$$f(B(\phi, y_0; h) = \sum_{\substack{\tau\in T\\|\tau|\geq 1}} \alpha(\tau)\beta(\tau)F(\tau)(y_0)\frac{h^{(|\tau|-1)}}{(|\tau|-1)!}$$

in which $\beta(\bullet) = 1$ and $\beta(\tau) = \prod_{i=1}^{m} \phi(\tau_i)$. Multiplying this by hm we get the following useful result:

Lemma 2 *Let $B(\phi, y_0; h)$ be a consistent B-series related to ODEs. Then*

$$hf(B(\phi, y_0; h) = B(\phi', y_0; h)$$

where

$$\phi'(\emptyset) = 0, \qquad \phi'(\bullet) = 1, \qquad \phi'(\tau) = |\tau| \prod_{i=1}^{m} \phi(\tau_i) \text{ if } \tau[\tau_1, \cdots, \tau_m].$$

Development of the B-series for the numerical solution

An s-stage Runge-Kutta method (RK), applied to the ODE (6) is given by

$$\begin{aligned} Y_i &= y_0 + h \sum_{j=1}^{s} a_{ij} f(Y_j), \qquad i = 1, \cdots, s, \\ y_1 &= y_0 + h \sum_{i=1}^{s} b_i f(Y_i) \end{aligned} \tag{8}$$

where $\{Y_i\}_{i=1}^s$ are the *internal stage values* and y_1 is the numerical solution after one step of stepsize h. The coefficients a_{ij} and b_i characterising the RK-method is given by the Butcher-tableau

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & & & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array} \qquad \text{or in matrix form as} \qquad \begin{array}{c|c} c & A \\ \hline & b^T \end{array}$$

We assume that the RK-nodes $c_i = \sum_{j=1}^s a_{ij}$ for $i = 1, \cdots, s$, or $c = A\mathbb{1}_s$ where $\mathbb{1} = [1, 1, \cdots, 1] \in \mathbb{R}^s$.

The next step is to write the numerical solution, given by (8) as a B-series related to the ODE.

$$\begin{aligned} Y_i = Y_i(h) &= \sum_{\tau \in T} \alpha(\tau) \varphi_i(\tau) F(\tau)(y_0) \frac{h^{|\tau|}}{|\tau|!}, \qquad i = 1, \cdots, s \\ y_1 = y_1(h) &= \sum_{\tau \in T} \alpha(\tau) \psi(\tau) F(\tau)(y_0) \frac{h^{|\tau|}}{|\tau|!}. \end{aligned}$$

Clearly, these series are consistent. If we insert this into (8) and apply Lemma 2 we get

$$\begin{aligned} B(\varphi_i, y_0; h) &= y_0 + \sum_{j=1}^{s} a_{ij} B(\varphi_j', y_0; h), \qquad i = 1, \cdots, s \\ B(\psi, y_0; h) &= y_0 + \sum_{i=1}^{s} b_i B(\varphi_i', y_0; h). \end{aligned}$$

Write out the series, compare term by term, and we get:

$$\varphi_i(\emptyset) = 0, \;\; \varphi_i(\tau) = \sum_{j=1}^{s} a_{ij}\varphi_j'(\tau), \qquad \psi(\emptyset) = 0, \;\; \psi(\tau) = \sum_{i=1}^{s} b_i \varphi_i'(\tau),$$

and we have shown:

Theorem 2 *The internal stage values Y_i, $i = 1, \cdots, s$ and the numerical solution y_1 given by* (8) *can be written as B-series related to ODEs, that is*

$$Y_i = B(\varphi_i, y_0; h), \qquad y_1 = B(\psi, y_0; h)$$

with coefficients

- $\varphi_i(\emptyset) = 1$, $\varphi(\bullet)_i = \sum_{j=1}^{s} a_{ij} = c_i$ *and if* $\tau = [\tau_1, \tau_2, \cdots, \tau_m] \in T$ *then*

  $\varphi_i(\tau) = |\tau| \sum_{j=1}^{s} a_{ij} \prod_{l=1}^{m} \varphi(\tau_l).$

- $\psi(\emptyset) = 1$, $\psi(\bullet) = \sum_{i=1}^{s} b_i$ *and if* $\tau = [\tau_1, \tau_2, \cdots, \tau_m] \in T$ *then*

  $\psi(\tau) = |\tau| \sum_{i=1}^{s} b_i \prod_{l=1}^{m} \varphi_i(\tau_l).$

Remark 1: In the literature, it is common to write the coefficients $\psi(\tau)$ as $\gamma(\tau)\bar{\psi}(\tau)$. Then $\gamma(\tau)$ is an integer, called the *density* of the tree t, while the *elementary weights* $\bar{\psi}(\tau)$ consists solely of the method coefficients. The order conditions is the written as $\bar{\psi}(\tau) = 1/\gamma(\tau)$.

The local truncation error satisfies

$$y(x_0 + h) - y_1(h) = \mathcal{O}(h^{p+1}) \;\text{ if }\; \psi(\tau) = 1, \quad \forall \tau \in T, \quad \rho(\tau) \leq p.$$

Thus we can conclude the section with the following result:

Theorem 3 *A Runge-Kutta method applied to an ODE is of order p if and only if*

$$\psi(\tau) = 1, \quad \forall \tau \in T, \quad \rho(\tau) \leq p.$$

The *if*-part of the theorem is clear from the discussion above. The *only if* comes from the fact that the elementary differentials are actually indepenent, see [1], pp. 146-147.

References

[1] J. C. Butcher. *Numerical methods for ordinary differential equations.* John Wiley & Sons Ltd., Chichester, 2003.

[2] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. I*, volume 8 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second revised edition, 2000. Nonstiff problems.

[3] J. D. Lambert. *Numerical methods for ordinary differential systems*. John Wiley & Sons Ltd., Chichester, 1991. The initial value problem.