to appear in: SIAM Journal of Numerical Analysis **34(5)** (1997)

SYMPLECTIC RUNGE-KUTTA SCHEMES I: ORDER CONDITIONS

M. SOFRONIOU* AND W. OEVEL†

Abstract. Much recent work has indicated that considerable benefit arises from the use of symplectic algorithms when numerically integrating Hamiltonian systems of differential equations. Runge-Kutta schemes are symplectic subject to a simple algebraic condition. Starting with Butcher's formalism it is shown that there exists a more natural basis for the set of necessary and sufficient order conditions for these methods; involving only $s(s+1)/2$ free parameters for a symplectic s stage scheme. A graph theoretical process for determining the new order conditions is outlined. Furthermore, it is shown that *any* rooted tree arising from the *same* free tree enforces the *same* algebraic constraint on the parametrised coefficients. When coupled with the standard simplifying assumptions for implicit schemes the number of order conditions may be further reduced. In the new framework a simple symmetry of the parameter matrix yields (not necessarily symplectic) self-adjoint methods. In this case the order conditions associated with even trees become redundant.

Key words. Runge-Kutta schemes, Hamiltonian systems, symplectic, trees, order conditions, symplectic integration, simplifying assumptions

AMS(MOS) subject classifications. primary 65L05; secondary 05C05

1 Introduction.

Consider the initial value problem for a system of first order differential equations:

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}) \qquad \mathbf{y}(t_0) = \mathbf{y}_0 \,. \tag{1.1}$$

*Wolfram Research Inc., 100 Trade Center Drive, Champaign, Illinois 61820, United States. Email: na.msofroniou@na-net.ornl.gov

†FB 17-Mathematik, Universität Paderborn, 33095 Paderborn, Germany. Email: walter@uni-paderborn.de

The Runge-Kutta method with s stages for the system is given by

$$\begin{aligned} \mathbf{k}_i &= \mathbf{f}(t_n + c_i h, \mathbf{y}_n + h \sum_{j=1}^{s} a_{ij}\mathbf{k}_j), \qquad i = 1, \ldots, s; \\ \mathbf{y}_{n+1} &= \mathbf{y}_n + h \sum_{j=1}^{s} b_j \mathbf{k}_j\,, \end{aligned} \tag{1.2}$$

where the parameters b_j and a_{ij} have to be chosen such that $\mathbf{y}_n$ represents an approximation of the Taylor expansion of the solution. The coefficients c_i are given by the row-sum conditions

$$c_i = \sum_{j=1}^{s} a_{ij}, \quad i = 1, \ldots, s. \tag{1.3}$$

They appear naturally in the derivation of order conditions for high order methods [10]. It has become customary to represent the free parameters of the method using a *Butcher tableau* [5]:

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array} \tag{1.4}$$

In the following we shall use the matrix $A = (a_{ij})$ and the column vector $\mathbf{b} = (b_1, \ldots, b_s)^T$ to denote the entries of such a tableau. We note that the Runge-Kutta map $\Psi_h : \mathbf{y}_n \mapsto \mathbf{y}_{n+1}$ is characterised by these parameters only up to permutation of the internal stages. Two tableaus $(A, \mathbf{b})$ and $(A', \mathbf{b}')$ yield the same map, if $A' = PAP^T$, $\mathbf{b}' = P\mathbf{b}$, where P is an arbitrary permutation matrix.

We briefly review some facts about the well-known Gauss-Legendre schemes which will be used later on. The parameters of these methods are defined by the set of *linear* equations [5, section 343]

$$\sum_{j=1}^{s} a_{ij}\, c_j^{k-1} = \frac{c_i^k}{k}\,, \quad \sum_{j=1}^{s} b_j\, c_j^{k-1} = \frac{1}{k}\,, \quad i, k = 1, \ldots, s$$

involving the Vandermonde matrix (c_j^{k-1}) generated by the roots $c_1, \ldots, c_s$ of the Legendre polynomial of degree s shifted to the interval $[0, 1]$. Different permutations of the parameters A and $\mathbf{b}$ arise from different orderings of these roots. We eliminate this freedom by requiring $0 < c_1 < \ldots < c_s < 1$. In this case

$$A = P(\mathbf{e}\mathbf{b}^T - A)P^T\,, \quad \mathbf{b} = P\mathbf{b} \tag{1.5}$$

with $\mathbf{e} = (1, \ldots, 1)^T$ and

$$P = \begin{pmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{pmatrix}. \tag{1.6}$$

This is easily derived from the symmetric distribution $c_{s+1-i} = 1 - c_i$ of the Legendre roots. The implication of this fact is as follows. The *adjoint* (or *reflection*) of a Runge-Kutta method $(A, \mathbf{b})$ is the scheme $(A', \mathbf{b}')$ whose map Ψ'_h provides the inverse of the map Ψ_h corresponding to the original scheme upon time reversal: $(\Psi'_{-h} \circ \Psi_h)(\mathbf{y}) = \mathbf{y}$. It is easily verified that the Runge-Kutta method $A' = P(\mathbf{e}\mathbf{b}^T - A)P^T$, $\mathbf{b}' = P\mathbf{b}$ with an arbitrary permutation matrix P defines the adjoint of the method $(A, \mathbf{b})$. Hence, due to (1.5), the Gauss-Legendre schemes are *self-adjoint* [5, theorem 344E]: they provide their own inverse. Self-adjoint methods are also sometimes referred to as *symmetric* methods[10, definition 8.7].

In the sequel, we will make use of the fact that Gauss-Legendre schemes of arbitrarily high order exist, if a sufficiently high number of stages is chosen: the Gauss-Legendre schemes have order $2s$ which is the maximal possible order for a Runge-Kutta method with s stages.

In recent years there has been much interest in the numerical treatment of Hamiltonian dynamical systems [19]. A *Hamiltonian system* of differential equations for the coordinates $\mathbf{y} = (\mathbf{p}, \mathbf{q})^T$ with $\mathbf{p}, \mathbf{q} \in \mathbf{R}^n$ is given by

$$\frac{d\mathbf{q}}{dt} = \frac{\partial H}{\partial \mathbf{p}}, \quad \frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{q}}$$

with some function $H \in \mathrm{C}^1(\mathbf{R}^{2n})$. In mechanics, $\mathbf{q}$ and $\mathbf{p}$ represent the generalised coordinates and momenta respectively and H is the energy of the system. A map from the phase space $\mathbf{R}^{2n}$ to itself is called *canonical* or *symplectic* if it leaves the differential form $\mathbf{dp} \wedge \mathbf{dq}$ invariant. An essential structural result for Hamiltonian systems is that the phase flow is symplectic [2]. A single step numerical method is called symplectic, if for any stepsize the corresponding map is symplectic, whenever the method is applied to a Hamiltonian system. The symplectic class of numerical methods recreate phase trajectories for Hamiltonian systems more faithfully than standard techniques since both the numerical as well as the exact flow share the crucial property of preservation of the symplectic form [29].

Around 1988, Lasagni [18], Sanz-Serna [26] and Suris [35] independently discovered that Runge-Kutta methods can be symplectic and the condition has since been proven to also be necessary. The map Ψ_h corresponding to a Butcher-tableau $(A, \mathbf{b})$ is symplectic for all Hamiltonian systems if and only if

$$b_i a_{ij} + b_j a_{ji} - b_i b_j = 0 \qquad i, j = 1, \ldots, s \tag{1.7}$$

or equivalently

$$M := BA + A^T B - \mathbf{b}\mathbf{b}^T = 0 \tag{1.8}$$

where

$$B = \mathrm{diag}(b_1, \ldots, b_s)\,.$$

We will refer to the constraints (1.7,1.8) as the *symplecticity condition.* An elegant proof, using differential forms, can be found in Sanz-Serna's recent survey article [27].

Here we will follow the formalism of Butcher who represents elementary differentials in terms of *rooted trees* [5]. There are several other approaches which deserve mention. The method of perturbed collocation has been used by Nørsett [21] and Wanner [22] in characterising Runge-Kutta methods and by Iserles [14] and Ramaswami [23] in characterising some symplectic Runge-Kutta methods. Collocation techniques can lead to a more direct method of construction than the elementary differential approach. However, order conditions are not available for all methods and therefore methods derived using these techniques form only a subset of possible Runge-Kutta methods. Albrecht has derived an alternative to Butcher's formalism which emphasises the similarities between Runge-Kutta and linear multi-step methods [1, 13]. Recently Munthe-Kaas has also presented an interesting approach based coordinate independent transformations using Lie series [20].

The symplecticity condition represents additional constraints on the free parameters which must be satisfied in conjunction with the order conditions. Although at first it may seem that for a given number of parameters some order conditions have to be sacrificed in favour of symplecticity, it turns out that a considerable number of order conditions become dependent as a consequence of (1.8). Roughly one equation per non-rooted remains [28]. Thus, the constraints imposed by symplecticity do not lead to order barriers but rather serve as simplifying assumptions. In fact, the (maximal) Gauss-Legendre schemes are symplectic [6, theorem 4.6].

It is known that partitioning via P-series can be used in a backward error analysis of numerical integrators [11]. The basic idea is to study the effect of a perturbation on the differential system. Here we will instead use a process of partitioning to study the effect of perturbation on the order conditions themselves. In the following we shall parametrise the Butcher matrix $A = (a_{ij})$ by another matrix $S = (s_{ij})$ according to

$$A = SB + \frac{1}{2}\,\mathbf{e}\mathbf{b}^T\,, \quad \text{i.e.} \quad a_{ij} = \left(s_{ij} + \frac{1}{2}\right) b_j\,, \quad i,j = 1,\ldots,s \tag{1.9}$$

and regard the scheme as being defined by the parameters $(S, \mathbf{b})$ instead of $(A, \mathbf{b})$. Some properties of methods are characterised in a simpler way with the new parameters S. In particular, symplecticity corresponds to skew-symmetry of S. The adjoint of the scheme $(S, \mathbf{b})$ is given by $(S', \mathbf{b}')$ with $S' = -PSP^T$, $\mathbf{b}' = P\mathbf{b}$, where P is an arbitrary permutation matrix. The most important issue in the following considerations is the reformulation of Butcher's order conditions for $(A, \mathbf{b})$ into new conditions for $(S, \mathbf{b})$ which reflect the structural properties of symplectic methods ($S = -S^T$) and self-adjoint methods ($S = -PSP^T$, $\mathbf{b} = P\mathbf{b}$) in a natural way. In these cases the new formulation reduces the number of parameters required to specify the scheme and simultaneously reduces the number of equations necessary to attain a certain order. For the symplectic case the reduction of rooted trees to non-superfluous non-rooted trees as found in [28] becomes trivial in the new formulation. For self-adjoint methods one finds a reduction from all rooted trees to rooted trees of odd order, since the new order equations associated with even trees turn out to be satisfied automatically. The standard simplifying assumptions for implicit Runge-Kutta schemes may also be modified and applied to reduce the number of order conditions still further [24]. In [25] it was shown that symplecticity serves to reduce the number of free parameters in the Butcher tableau (1.4). It is believed that the approach outlined here is more direct and hence more constructive. We will give a simple application as justification of this claim.

In section 2 we review the necessary graph theoretical definitions. Most are variants of those found in [5]. We assume, where appropriate, some of the more technical definitions given in [28]. Butcher's formalism is reviewed in which the order condition are formulated in terms of weights and densities associated with rooted trees. In section 3 the Butcher matrix is parametrised according to (1.9) and the characterisation of symplectic and adjoint schemes is discussed. In section 4 Butcher's order conditions are reformulated as a set of *homogeneous* polynomial equations in the new parameters $(S, \mathbf{b})$. This requires the introduction of a new weight function and a corresponding analogue of (the reciprocal of) density for rooted trees which are combinations of the corresponding objects in Butcher's classical order theory. The new weight has the property of being invariant under root-shifting (up to a change of sign), when the scheme is symplectic. Thus, in the symplectic case, the order conditions are given by trees rather than rooted trees, since the equations do not depend on the location of the root. This was the main result of Sanz-Serna and Abia in [28]. The present formulation seems more natural since it resolves the distinction between the homogeneous and the inhomogeneous form of the order conditions as given by Sanz-Serna and Abia. Imposing a simple symmetry on the parameter matrix S, corresponding to self-adjoint methods, the new order conditions are automatically satisfied for all rooted trees of even order. In section 5 a direct construction of the new density for a certain class of trees is given. Generating functions are derived, from which the values of these densities can be obtained for all trees relevant to derive symplectic Runge-Kutta methods of order $p \leq 10$ with stages $s \leq 5$. In section 6 Butcher's simplifying assumptions are modified in order to yield simplifications of the new order conditions. In section 7 a simple application of the re-parametrised order conditions is given with an explicit proof for the non-existence of 2-stage 3rd order Runge-Kutta methods with real spectra.

2 Butcher's formalism.

We review some of the standard notions of graph theory relevant in Butcher's formulation of the order conditions.

i. A *labeled tree* $\lambda\tau$ is a labeled graph of vertices V and edges E, denoted (V, E), such that for any pair of distinct vertices $v, w \in V$ there exists a unique path joining v and w.

ii. A *rooted labeled tree* $\rho\lambda\tau$ is a labeled tree where one of the vertices, the root, has been highlighted. The vertices adjacent to the root are termed the *children* of the root - the remaining children are defined recursively. This can be thought of as associating with each edge a *direction* away from the root through all parent-child pairs. Graphically we highlight the root by encircling it [12].

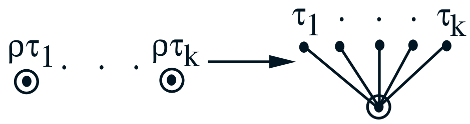


Figure 1: Grafting the trees $\rho\tau_1, \ldots, \rho\tau_k$, onto a new root yields the tree $[\rho\tau_1 \cdots \rho\tau_k]$.

iii. A *rooted tree* $\rho\tau$ is the equivalence class of rooted labeled trees isomorphic under relabeling of the vertices.

iv. Let $r(\rho\tau)$ denote the *order* or *number of vertices* of a tree. We will occasionally denote a tree of order r by an r-tree.

v. For each rooted tree $\rho\tau$ we denote by $\gamma(\rho\tau)$ the *density* of the tree which can be thought of as a measure of non-bushiness. It is defined recursively as

$$\gamma([\rho\tau_1 \cdots \rho\tau_k]) = r([\rho\tau_1 \cdots \rho\tau_k]) \prod_{i=1}^{k} \gamma(\rho\tau_i) \tag{2.1}$$

with $\gamma(\rho\tau) = 1$ if $r(\rho\tau) = 1$. The rooted tree $[\rho\tau_1 \cdots \rho\tau_k]$ is obtained by grafting the original trees $\rho\tau_i$ onto a new root [5], as in figure 1. An example for the calculation of the density γ is given in figure 2. We denote the *reciprocal of density* by

$$\delta(\rho\tau) = \frac{1}{\gamma(\rho\tau)} \; .$$

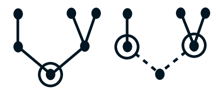
The elementary differentials occurring in the Taylor expansion of the solution of a dynamical system are in one-to-one correspondence with rooted trees $\rho\tau$ [5, theorem 305C]. The coefficients of the expansion involve the values $\delta(\rho\tau)$. For a Runge-Kutta method (1.2), the expansion of the map $\mathbf{y}_n \to \mathbf{y}_{n+1}$ into powers of h reproduces the elementary differentials with coefficients, which involve polynomial expressions of the free parameters b_i and a_{ij} of the method. These polynomials are called the *weights* $\Phi(\rho\tau)$ and may be defined via the *weight vectors* $\varphi = (\varphi_1, \ldots, \varphi_s)$ which are given recursively by

$$\varphi_i([\rho\tau_1 \cdots \rho\tau_k]) = \prod_{m=1}^{k} \sum_{j=1}^{s} a_{ij} \varphi_j(\rho\tau_m) \tag{2.2}$$

with $\varphi_i(\rho\tau) = 1$ if $r(\rho\tau) = 1$. They give rise to the weight

$$\Phi(\rho\tau) \;=\; < \mathbf{b}, \varphi(\rho\tau) > \;=\; \sum_{i=1}^{s} b_i \, \varphi_i(\rho\tau) \; . \tag{2.3}$$

One can visualise the weight graphically in the following way. Choose a labeling of the rooted r-tree given by the vertex set $V = \{1, \ldots, r\}$ and the set of edges $E \subset V \times V$. The



$$\gamma(\rho\tau) = 6 \cdot 2 \cdot 3 = 36$$

Figure 2: The density of a rooted tree $\rho\tau$, $r(\rho\tau) = 6$.

ordering of the edge $(m, n) \in E$ represents the parent/child relation induced by the root. The vector $\mathbf{b}$ is attached to the root, with each edge we associate a copy of the matrix A. Then a direct definition of the weight function is given by

$$\Phi(\rho\tau) = \sum_{j_1=1}^{s} \cdots \sum_{j_r=1}^{s} b_{j_\rho} \prod_{(m,n)\in E} a_{j_m j_n} \tag{2.4}$$

where $\rho \in V$ is the label of the root. This definition is invariant under re-labeling so that the weight is associated with the equivalence class $\rho\tau$ rather than with a labeled representative $\rho\lambda\tau$. These weights are clearly linear functions in the parameters b_i and homogeneous expressions in a_{ij} where the degree of the polynomial coincides with the number of edges in the tree. For a method (1.2) to attain order $p \geq 1$ necessary and sufficient conditions are provided by the equations [5, theorem 307B]

$$\Phi(\rho\tau) = \frac{1}{\gamma(\rho\tau)} = \delta(\rho\tau) \quad \text{for each rooted tree } \rho\tau \text{ of order } r(\rho\tau) \leq p\ . \tag{2.5}$$

For generic values of the parameters any collection of weights is functionally independent if a sufficiently large number of stages s is chosen.

However, certain points exist in the parameter space, where the rank of these equations is less than maximal. Indeed, Sanz-Serna and Abia [28] showed that due to the symplecticity condition (1.8) many of the equations no longer remain independent. In particular, we will show that the weights may be redefined in such a way that, for symplectic schemes, the new weights change only in sign when the root of the tree is shifted. Consequently, for symplectic schemes, it suffices to impose (approximately) as many order conditions as the number of non-rooted trees with p or less vertices.

3 Parametrisation of the Butcher matrix.

For the symplectic case the free parameters $(A, \mathbf{b})$ in the Runge-Kutta tableau (1.4) are to be determined by the order conditions (2.5) as well as the symplecticity constraint (1.7). Instead of using the set of s^2 parameters a_{ij} one may parametrise the Butcher matrix by the $s \times (s-1)/2$ entries of a skew symmetric matrix S in such a way that the symplecticity condition is satisfied identically. Together with the parameters b_i one finds a reduced set of $s \times (s+1)/2$ free parameters $(S, \mathbf{b})$ which are to be determined *solely* from the order conditions.

Lemma 3.1 [8, theorem 4.4] *With the parametrisation (1.9) a Runge-Kutta scheme is symplectic if S is skew-symmetric. For non-degenerate B skew-symmetry of S is also necessary for symplecticity.*

Proof Insertion of the parametrisation (1.9) into the symplecticity condition (1.8) yields $M = B(S + S^T)B = 0$. □

We note that for symplectic schemes one may assume that the weights $\mathbf{b}$ are indeed non-zero: if $b_j = 0$ for some j then (1.7) implies $b_i\, a_{ij} = 0$. Therefore the jth stage does not contribute to the quadrature and the method reduces to one with fewer stages. Symplectic Runge-Kutta methods are necessarily implicit since the diagonal elements $a_{ii} = b_i/2$ do not vanish. The choice $s_{ij} = -1/2$ for $j > i$ leads to the lower triangular form

$$\begin{array}{c|ccccc} & b_1/2 & 0 & \cdots & \cdots & 0 \\ & b_1 & b_2/2 & \ddots & \ddots & \vdots \\ & b_1 & b_2 & b_3/2 & \ddots & \vdots \\ & \vdots & \vdots & \vdots & \ddots & \vdots \\ & b_1 & b_2 & b_3 & \cdots & b_s/2 \\ \hline & b_1 & b_2 & b_3 & \cdots & b_s \end{array}$$

i.e. to diagonally implicit symplectic Runge-Kutta methods [28].

Lemma 3.2 *With the parametrisation (1.9) the adjoint (reflected) method of the Runge-Kutta scheme $(S, \mathbf{b})$ is given by $(S', \mathbf{b}')$ with $S' = -PSP^T$, $\mathbf{b}' = P\,\mathbf{b}$, where P is an arbitrary permutation matrix. Consequently, methods satisfying*

$$S = -PSP^T\,, \quad \mathbf{b} = P\mathbf{b} \tag{3.1}$$

with some permutation matrix P are self-adjoint.

Proof The adjoint of the method $(A, \mathbf{b})$ is given by $(A', \mathbf{b}')$ with $A' = P(\mathbf{e}\mathbf{b}^T - A)P^T$, $\mathbf{b}' = P\mathbf{b}$. Insertion of the parametrisations

$$A = SB + \frac{1}{2}\,\mathbf{e}\mathbf{b}^T\,, \quad A' = S'B' + \frac{1}{2}\,\mathbf{e}\mathbf{b}'^T \quad \text{with} \quad \mathbf{b} = B\mathbf{e}\,, \quad \mathbf{b}' = B'\mathbf{e}$$

yields

$$(S' + PSP^T)B' = P\Big(\,\frac{1}{2}\,\mathbf{e}\mathbf{e}^T - S\,\Big)P^T\,(PBP^T - B')\,.$$

This equation is satisfied by $S' = -PSP^T$ and $\mathbf{b}' = P\mathbf{b}$ (equivalent to $B' = PBP^T$). By uniqueness, the Runge-Kutta map $\mathbf{y}_n \mapsto \mathbf{y}_{n+1}$ generated by $(S', \mathbf{b}')$ yields *the* reflection of the map generated by $(S, \mathbf{b})$. □

We note that amongst all Runge-Kutta schemes the sub-classes of symplectic and self-adjoint methods are characterised by very simple symmetries of the matrix S. Concrete methods may be constructed by starting with an appropriate S which is parametrised by a reduced set of parameters such that the required symmetry is satisfied. In the search for suitable values of these parameters one needs to satisfy the order equations associated with Runge-Kutta schemes. In the next section we reformulate Butcher's classical order conditions so that for symplectic and self-adjoint schemes these equations are drastically simplified, yielding a reduced set of constraints on the reduced set of parameters.

4 Transforming the order conditions.

We examine the effect of the parametrisation (1.9) with arbitrary (not necessarily skew-symmetric) S on the order conditions and re-derive these conditions in terms of the new parameters $(S, \mathbf{b})$. Of course, the order conditions (2.5) immediately lead to equations in terms of $(S, \mathbf{b})$ by inserting (1.9) into the weight function Φ. However, in terms of the original parameters (a_{ij}) the weights are *homogeneous* polynomials, which now turn into *inhomogeneous* polynomials of the new parameters s_{ij}. Our aim is to combine the old weights associated with different rooted trees into a new weight function which is homogeneous in the new unknowns. The easiest way to do this is to take the homogeneous expressions Φ and simply replace every occurrence of a_{ij} by $s_{ij}\, b_j$. This defines a different weight $\tilde{\Phi}$, since the term $b_j/2$ is omitted in the substitution $a_{ij} \to s_{ij} b_j$. A direct construction of $\tilde{\Phi}$, used in the following definition 4.1, is obtained from (2.4). Recursive relations may be derived from which it becomes apparent that for any rooted tree the new weight is just a combination of Butcher's old weights for a collection of other trees. It turns out that the new weight reflects the effect of symplecticity in a simple way: it changes only its sign under root-shifting.

Definition 4.1 (Analogue of weight) *For any rooted tree $\rho\tau$, define the new weight function $\tilde{\Phi}(\rho\tau)$ as $\Phi(\rho\tau)$ with a_{ij} replaced by $s_{ij} b_j$, i.e. (using a labeling)*

$$\tilde{\Phi}(\rho\tau) \;=\; \sum_{j_1=1}^{s} \cdots \sum_{j_r=1}^{s} b_{j_\rho} \prod_{(m,n)\in E} s_{j_m j_n} b_{j_n}\,, \qquad r = r(\rho\tau)\,. \tag{4.1}$$

The degree of the polynomial in $\mathbf{b}$ and S is given by the number r of vertices and by the number $r-1$ of edges, respectively. Graphically (4.1) may be visualised by associating the vector $\mathbf{b}$ with each vertex and the matrix S with each edge. It remains to define a suitable analogue $\tilde{\delta}(\rho\tau)$ of the reciprocal of density $\delta(\rho\tau)$, such that the order conditions (2.5) are transformed into $\tilde{\Phi}(\rho\tau) = \tilde{\delta}(\rho\tau)$. To this end we first establish a relation between the weight functions $\tilde{\Phi}$ and Φ and then use this relation to define $\tilde{\delta}$ in terms of δ.

The arguments become more transparent when the new quantities are regarded as deformations of the old ones. We construct a weight $\Phi^{(\epsilon)}$ via (2.2,2.3,2.4) but with a_{ij} replaced by

$$a_{ij}^{(\epsilon)} = a_{ij} + \epsilon\, b_j \ , \ \text{i.e. } \ A^{(\epsilon)} = (a_{ij}^{(\epsilon)}) = A + \epsilon\, \mathbf{e}\, \mathbf{b}^T \ ,$$

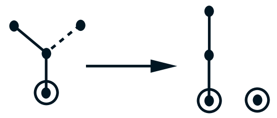


Figure 3: Chopping an edge (dashed) from a 4-tree and the resulting 1-partition.

where ϵ is an arbitrary parameter. For the special value $\epsilon = -1/2$ one finds $a_{ij}^{(-1/2)} = s_{ij} b_j$, so that $\tilde{\Phi}(\rho\tau) = \Phi^{(-1/2)}(\rho\tau)$ coincides with the weight of definition 4.1. In order to establish the relation between $\Phi^{(\epsilon)}$ and Φ we expand the perturbed weight as a power series in ϵ. We need the following definition.

Definition 4.2 (Chopping and Partitioning) Chopping *is the process of removing edges from a rooted tree $\rho\tau$. In the removal of each edge the sub-tree containing the original root maintains its root. The remaining sub-tree is rooted at the vertex where the edge was chopped. A k-*partition *of a rooted tree $\rho\tau$ is the collection of $k+1$ rooted sub-trees resulting from chopping k edges. We denote a k-partition by P_k.*

An illustration of the processes of chopping and partitioning is given in figure 3. Upon replacing A by $A^{(\epsilon)}$ in (2.4) one obtains $\Phi^{(\epsilon)}$ as a polynomial in ϵ. The power ϵ^p arises by choosing p edges, for which $a_{j_m j_n}$ is replaced by $\epsilon\, b_{j_n}$, whereas the original matrix A is used for all the remaining edges. In (2.4) the product over all edges factorises and yields the product of the weights Φ associated with the rooted sub-trees arising from the k-partition. Hence, one finds

$$\begin{aligned} \Phi^{(\epsilon)}(\rho\tau) &= \Phi(\rho\tau) + \epsilon \sum_{P_1} \Phi(\rho\tau_0)\Phi(\rho\tau_1) + \epsilon^2 \sum_{P_2} \Phi(\rho\tau_0)\Phi(\rho\tau_1)\Phi(\rho\tau_2) \\ &\quad + \cdots + \epsilon^{r-1} \Big(\sum_{j=1}^{s} b_j\Big)^r , \end{aligned} \tag{4.2}$$

where $r = r(\rho\tau)$ is the order of the tree. The sum associated with ϵ^k extends over all k-partitions of $\rho\tau$ (i.e. over all $\binom{r-1}{k}$ subsets of E with k elements), where each induces the decomposition into the $k+1$ sub-trees $\rho\tau_0, \ldots, \rho\tau_k$. The last term in (4.2) corresponds to chopping all edges, so that the tree decomposes into r isolated roots, each of which gives rise to the weight

$$\Phi(\rho\tau) = \Phi^{(\epsilon)}(\rho\tau) = \sum_{j=1}^{s} b_j \quad \text{for} \quad r(\rho\tau) = 1 . \tag{4.3}$$

An alternative recursive formula is obtained by performing the same perturbation starting with $\Phi^{(\epsilon)}$. With $a_{ij}^{(\epsilon+\epsilon')} = a_{ij}^{(\epsilon)} + \epsilon'\, b_j$ one expands $\Phi^{(\epsilon+\epsilon')}$ into powers of ϵ' around $\Phi^{(\epsilon)}$ and obtains

$$\begin{aligned} \Phi^{(\epsilon+\epsilon')}(\rho\tau) &= \Phi^{(\epsilon)}(\rho\tau) + \epsilon' \sum_{P_1} \Phi^{(\epsilon)}(\rho\tau_0)\Phi^{(\epsilon)}(\rho\tau_1) + \epsilon'^2 \sum_{P_2} \Phi^{(\epsilon)}(\rho\tau_0)\Phi^{(\epsilon)}(\rho\tau_1)\Phi^{(\epsilon)}(\rho\tau_2) \\ &\quad + \cdots + \epsilon'^{r-1} \Big(\sum_{j=1}^{s} b_j\Big)^r , \end{aligned}$$

analogous to (4.2). Choosing $\epsilon' = -\epsilon$ leads to $\Phi^{(0)} = \Phi$ in the left hand side, so that solving for $\Phi^{(\epsilon)}(\rho\tau)$ one obtains

$$\begin{aligned}\Phi^{(\epsilon)}(\rho\tau) &= \Phi(\rho\tau) + \epsilon \sum_{P_1} \Phi^{(\epsilon)}(\rho\tau_0)\Phi^{(\epsilon)}(\rho\tau_1) - \epsilon^2 \sum_{P_2} \Phi^{(\epsilon)}(\rho\tau_0)\Phi^{(\epsilon)}(\rho\tau_1)\Phi^{(\epsilon)}(\rho\tau_2) \\ &\qquad \pm \cdots + (-1)^r \epsilon^{r-1} \Big(\sum_{j=1}^{s} b_j\Big)^r .\end{aligned} \tag{4.4}$$

Coupled with (4.3) this serves as an alternative recursive definition for $\Phi^{(\epsilon)}$, since the right hand side involves only sub-trees of lower orders.

We modify the reciprocal of density, δ, by duplicating the relations (4.2,4.4), setting $\epsilon = -1/2$ and replacing the weight by the density.

Definition 4.3 (Analogue of reciprocal of density) *For any rooted tree $\rho\tau$ of order $r = r(\rho\tau)$ we define the new density*[1] $\tilde{\delta}(\rho\tau)$ *as*

$$\tilde{\delta}(\rho\tau) = \delta(\rho\tau) - \frac{1}{2}\sum_{P_1} \delta(\rho\tau_0)\delta(\rho\tau_1) + \Big(\frac{1}{2}\Big)^2 \sum_{P_2} \delta(\rho\tau_0)\delta(\rho\tau_1)\delta(\rho\tau_2) \mp \cdots + \Big(-\frac{1}{2}\Big)^{r-1}$$

or equivalently, by

$$\tilde{\delta}(\rho\tau) = \delta(\rho\tau) - \frac{1}{2}\sum_{P_1} \tilde{\delta}(\rho\tau_0)\tilde{\delta}(\rho\tau_1) - \Big(\frac{1}{2}\Big)^2 \sum_{P_2} \tilde{\delta}(\rho\tau_0)\tilde{\delta}(\rho\tau_1)\tilde{\delta}(\rho\tau_2) - \cdots - \Big(\frac{1}{2}\Big)^{r-1} .$$

For the trivial tree with order $r = 1$, $\tilde{\delta}(\rho\tau) = \delta(\rho\tau) = 1$.

From a graph theoretical point of view it is not at all clear that these two alternative definitions are equivalent. However, in the present context of order conditions remark 4.5, given later in this section, yields a suitable argument to establish equivalence.

By construction, the modified weights and reciprocals of density coincide with Butcher's original versions up to lower order rooted trees. By induction Butcher's order conditions (2.5) immediately translate to the following theorem.

Theorem 4.4 (New Order Conditions) *A Runge-Kutta method is of order p if and only if $\tilde{\Phi}(\rho\tau) = \tilde{\delta}(\rho\tau)$ for all rooted trees of order $r(\rho\tau) \leq p$.*

Proof The inductive step from $p-1$ to p is given as follows. For any rooted tree $\rho\tau$ of order p one has

$$\Big(\tilde{\Phi}(\rho\tau) - \tilde{\delta}(\rho\tau)\Big) - \Big(\Phi(\rho\tau) - \delta(\rho\tau)\Big) =$$
$$\sum_{k=1}^{p-1} \Big(-\frac{1}{2}\Big)^k \sum_{P_k} \Big(\Phi(\rho\tau_0)\cdots\Phi(\rho\tau_k) - \delta(\rho\tau_0)\cdots\delta(\rho\tau_k)\Big) = 0 ,$$

since the k-partitions involve rooted trees of maximal order $p-1$, for which $\Phi(\rho\tau_j) = \delta(\rho\tau_j)$ is assumed. Hence, for all rooted trees of order p the modified order conditions are equivalent to the old ones. □

[1] We will refer to $\tilde{\delta}$ as a density, despite the fact that it may attain non-positive values. We note that one has to modify δ rather than γ, since for many trees $\tilde{\delta} = 0$.

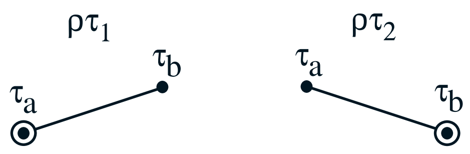


Figure 4: $\rho\tau_1$ and $\rho\tau_2$ can be considered as joining of two lower order rooted trees $\rho\tau_a$ and $\rho\tau_b$ by their roots, maintaining one of the original roots.

Remark 4.5 Based on theorem 4.4 we may formulate the following general principle which will be useful to derive various properties of the new density:

Any relation satisfied by the weight functions $\Phi, \tilde{\Phi}$ is also satisfied by the densities $\delta, \tilde{\delta}$.

We note that Runge-Kutta schemes of arbitrarily high order exist. For a relation between Φ and $\tilde{\Phi}$ involving a finite set of rooted trees we consider a scheme of sufficiently high order such that Φ and $\tilde{\Phi}$ attain the numerical values δ and $\tilde{\delta}$, respectively, for all trees involved. The resulting relation between δ and $\tilde{\delta}$ is a purely graph theoretical statement. Application of this principle to the relation (4.4) with $\epsilon = -1/2$ immediately grants the equivalence of the two recurrencies for $\tilde{\delta}$ given in definition 4.3.

In [28] it was shown that in the symplectic case the order conditions are related under root-shifting. A simple argument for this fact is given by the following property of the new weight and the new density. Consider two rooted trees $\rho\tau_1$ and $\rho\tau_2$ arising from the same unrooted tree τ with the roots one edge apart (see figure 4).

Lemma 4.6 *The new density changes its sign under root-shifting: $\tilde{\delta}(\rho\tau_1) = -\tilde{\delta}(\rho\tau_2)$. For symplectic schemes the new weight changes its sign under root-shifting: $\tilde{\Phi}(\rho\tau_1) = -\tilde{\Phi}(\rho\tau_2)$.*

Proof Consider

$$\tilde{\Phi}(\rho\tau_1) = \sum_{i,j=1}^{s} b_i b_j s_{ij} \tilde{\varphi}_i(\rho\tau_a)\tilde{\varphi}_j(\rho\tau_b)$$

where $\tilde{\varphi}$ represents the modified weight vector (2.2) with a_{ij} replaced by $s_{ij}b_j$. Due to the skew symmetry of S for symplectic schemes this weight changes its sign when τ_a and τ_b are interchanged. Since symplectic schemes of arbitrarily high order exist, we may invoke the principle formulated in remark 4.5 and replace $\tilde{\Phi}$ by $\tilde{\delta}$. $\square$

A direct consequence of this property is

$$\tilde{\Phi}(\rho\tau_1) - \tilde{\delta}(\rho\tau_1) \;=\; -\left(\tilde{\Phi}(\rho\tau_2) - \tilde{\delta}(\rho\tau_2)\right)$$

for symplectic schemes which gives us the following corollary:

Corollary 4.7 (Equivalence) *For symplectic schemes the order conditions for two root-shifted trees coincide.*

The significance of this corollary is the following. For symplectic schemes we need only consider one order condition per (non-rooted) tree, as the corresponding constraints on the parameters coincide for all rooted representatives. This means a significant reduction in the number of equations which must be satisfied. This result, already given in [28], also carries over into the new framework.

Further reductions arise from *superfluous* trees. These are constructed by joining two copies of a lower order rooted tree by their roots. Their order is necessarily even. It was pointed out in [28] that the order conditions for superfluous trees are implied by imposing lower order trees. In the new formulation we have the stronger statement that these conditions are redundant.

Corollary 4.8 (Redundancy) *Let $\rho\tau$ be superfluous. Then $\tilde{\delta}(\rho\tau) = 0$ and, for symplectic schemes, $\tilde{\Phi}(\rho\tau) = 0$.*

Proof Consider the rooted trees $\rho\tau_1$ and $\rho\tau_2$ depicted in figure 4 with $\tau_a = \tau_b$. Shifting the root by an edge $\tilde{\delta}$ and, assuming symplecticity, $\tilde{\Phi}$ changes in sign. Since root-shifting reproduces the superfluous rooted tree, both $\tilde{\delta}$ and $\tilde{\Phi}$ must vanish. □

We have demonstrated that the new order conditions inherit a simple root-shifting structure from the symplectity requirement, so that they seem more appropriate for symplectic schemes. In the following we will show that the new order conditions also inherit some simplified structure for self-adjoint schemes.

Lemma 4.9 (New weights for adjoint schemes) *Let $\tilde{\Phi}$ and $\tilde{\Phi}'$ be the new weights associated with the Runge-Kutta scheme $(S, \mathbf{b})$ and its adjoint scheme $(S', \mathbf{b}')$, respectively, represented by $S' = -PSP^T$, $\mathbf{b}' = P\,\mathbf{b}$ with some permutation matrix P. Then*

$$\tilde{\Phi}'(\rho\tau) = (-1)^{r-1}\,\tilde{\Phi}(\rho\tau) \quad \textit{with} \quad r = r(\rho\tau)\,.$$

Proof Let $p : \{1, \ldots, s\} \mapsto \{1, \ldots, s\}$ be the permutation given by P. With

$$S' = (s'_{ij}) = -PSP^T = (-s_{p(i),p(j)}) \quad \text{and} \quad \mathbf{b}' = (b'_j) = P\mathbf{b} = (b_{p(j)})$$

the resummations $j_k : 1, \ldots, s \to j_k : p(1), \ldots, p(s)$ in (4.1) yield

$$\begin{aligned} \tilde{\Phi}(\rho\tau) &= \sum_{j_1=1}^{s} \cdots \sum_{j_r=1}^{s} b_{p(j_\rho)} \prod_{(m,n)\in E} s_{p(j_m)p(j_n)} b_{p(j_n)} \\ &= \sum_{j_1=1}^{s} \cdots \sum_{j_r=1}^{s} b'_{j_\rho} \prod_{(m,n)\in E} (-s'_{j_m j_n}) b_{j_n} \;=\; (-1)^{r-1}\,\tilde{\Phi}'(\rho\tau)\,. \end{aligned}$$

□

For self-adjoint schemes this yields $\tilde{\Phi}(\rho\tau) = -\tilde{\Phi}(\rho\tau)$ if $r(\rho\tau)$ is even:

Corollary 4.10 (New weights for self-adjoint schemes) *For self-adjoint Runge-Kutta schemes (3.1) the new weight vanishes for all rooted trees $\rho\tau$ of even order: $\tilde{\Phi}(\rho\tau) = 0$.*

Since self-adjoint schemes of arbitrarily high order exist, one may invoke the principle of remark 4.5 and replace $\tilde{\Phi}$ by $\tilde{\delta}$:

Corollary 4.11 (New densities for even order trees) *The new density vanishes for all rooted trees $\rho\tau$ of even order: $\tilde{\delta}(\rho\tau) = 0$.*

Combining these results one finds:

Corollary 4.12 (New order conditions for adjoint and self-adjoint schemes) *The new order conditions for a scheme coincide with those for the adjoint scheme. For self-adjoint schemes the new order conditions arising from rooted trees of even order are automatically satisfied.*

Proof According to lemma 4.9 one has $\tilde{\Phi}'(\rho\tau) - \tilde{\delta}(\rho\tau) = \tilde{\Phi}(\rho\tau) - \tilde{\delta}(\rho\tau)$ for rooted trees $\rho\tau$ of odd order. For even trees the order condition is $\tilde{\Phi}'(\rho\tau) = -\tilde{\Phi}(\rho\tau) = 0$, since $\tilde{\delta}(\rho\tau) = 0$ according to corollary 4.11. For self-adjoint schemes $\tilde{\Phi}(\rho\tau) = 0 = \tilde{\delta}(\rho\tau)$ follows from corollary 4.10. □

Corollary 4.12 implies that self-adjoint Runge-Kutta schemes are necessarily of even order (cf. [10, theorem 8.10]). It is also known that a collocation method with so-called symmetrically distributed collocation points, $c_i = 1 - c_{s+1-i}$, is self-adjoint [10, theorem 8.9]. In the new formalism self-adjoint methods possess collocation points symmetrically distributed about the origin; $\tilde{c}_i = -\tilde{c}_{s+1-i}$ with $\tilde{c}_i$ denoting the modified row-sum conditions (see (6.1)).

In the search for self-adjoint schemes one may parametrise S in such a way that (3.1) is satisfied identically for some prescribed permutation. A good choice of the permutation $P \mathrel{\hat{=}} p : \{1, \ldots, s\} \mapsto \{1, \ldots, s\}$ minimizing the reduction of free parameters is as follows. Let $\{(i, j), (p(i), p(j)), (p(p(i)), p(p(j)), \ldots\}$ be the orbit of some index pair $i, j \in \{1, \ldots, s\}$ under repeated application of the permutation. Note that for self-adjoint schemes all entries of the matrix $S = \pm P \cdots PSP^T \cdots P^T$ corresponding to this orbit coincide with s_{ij} up to a change in sign. Hence the length of these orbits should be as small as possible. Further, it should be even, since otherwise $s_{ij} = 0$ and this entry would be lost as a free parameter. Hence, a permutation P satisfying $P^2 = 1$ and having a minimal number of fixed points gives rise to self-adjoint schemes without implying an exorbitant loss of free parameters. In particular, P given by (1.6) meets these requirements.

Order	1	2	3	4	5	6	7	8	9	10
General RK	1	2	4	8	17	37	85	200	486	1205
Symplectic RK	1	1	2	3	6	10	21	40	87	184
Symplectic self-adjoint RK	1	1	2	2	5	5	16	16	63	63

Table 1: Number of rooted trees required for Runge-Kutta methods of order $p \leq 10$.

We finally combine symplecticity and self-adjointness induced by the permutation (1.6). *Persymmetric* matrices

$$S = \begin{pmatrix} 0 & -s_{21} & -s_{31} & \cdots & \cdots & \\ s_{21} & 0 & -s_{32} & \cdots & & \cdots \\ s_{31} & s_{32} & \ddots & & \cdots & \cdots \\ \cdots & \cdots & & \ddots & -s_{32} & -s_{31} \\ \cdots & & \cdots & s_{32} & 0 & -s_{21} \\ & \cdots & \cdots & s_{31} & s_{21} & 0 \end{pmatrix}$$

together with $\mathbf{b} = (b_1, b_2, \cdots, b_2, b_1)^T$ yield symplectic self-adjoint schemes. The number of free parameters can be expressed as $\lfloor (s+1)^2/4 \rfloor$ where $\lfloor x \rfloor$, $x \in \mathbb{Q}$ denotes the integer part of x. This is a reduction by approximately a factor of 4 over the number of parameters available for a general s-stage method. Due to the previous results the number of order conditons are reduced to those generated by non-rooted trees of odd order:

Corollary 4.13 *A symplectic self-adjoint Runge-Kutta scheme is of order p, if and only if the order conditions are satisfied for all odd non-rooted trees with up to p vertices. The order of such methods is necessarily even.*

We remark that this result was derived using the new order conditions $\tilde{\Phi}(\rho\tau) = \tilde{\delta}(\rho\tau)$. Of course, since the new quantities correspond to Butcher's original versions up to lower order trees, corollary 4.13 also holds for the equations $\Phi(\rho\tau) = \delta(\rho\tau)$, which only need to be imposed for one rooted representative of odd trees with up to p vertices.

The reduction in the number of trees necessary for order $p \leq 10$ is given in table 1 (see [28]). Due to root-shifting properties in the symplectic case only one rooted representative is needed for any non-superfluous non-rooted tree, whereas in the general case all rooted trees of order $r \leq p$ are necessary. In the symplectic self-adjoint case no trees of even order need to be considered. In table 2 we give sufficient rooted trees and related properties for symplectic schemes to obtain order $p \leq 10$. The rooted trees $\rho\tau_1, \ldots, \rho\tau_{64}$ represent all non-superfluous trees of orders $p \leq 6$ and hence determine all order conditions independent of the number of stages. Many more trees of order $r > 6$ exist

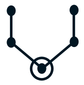
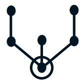
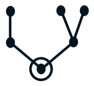
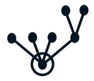
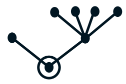
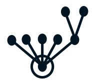
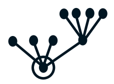
	$\rho\tau$	$r(\rho\tau)$	$\alpha(\rho\tau)$	$\gamma(\rho\tau)$	$\delta(\rho\tau)$	$\tilde{\delta}(\rho\tau)$	$\sum_{\rho\tau\in\tau}\beta(\rho\tau)$	$\bar{\beta}(\rho\tau)$
$\rho\tau_1$		1	1	1	1	1	1	1
$\rho\tau_3$		3	1	3	$\frac{1}{3}$	$\frac{1}{12}$	3	3
$\rho\tau_4$		4	1	4	$\frac{1}{4}$	0	4	4
$\rho\tau_{51}$		5	1	5	$\frac{1}{5}$	$\frac{1}{80}$	5	5
$\rho\tau_{52}$		5	4	15	$\frac{1}{15}$	$\frac{1}{240}$	60	60
$\rho\tau_{53}$		5	3	20	$\frac{1}{20}$	$\frac{1}{120}$	60	60
$\rho\tau_{61}$		6	1	6	$\frac{1}{6}$	0	6	6
$\rho\tau_{62}$		6	10	12	$\frac{1}{12}$	0	120	120
$\rho\tau_{63}$		6	15	24	$\frac{1}{24}$	0	360	360
$\rho\tau_{64}$		6	10	36	$\frac{1}{36}$	0	360	360
$\rho\tau_{71}$		7	1	7	$\frac{1}{7}$	$\frac{1}{448}$	7	7
$\rho\tau_{72}$		7	20	21	$\frac{1}{21}$	$\frac{1}{1344}$	420	420
$\rho\tau_{73}$		7	6	35	$\frac{1}{35}$	$\frac{1}{2240}$	210	210
$\rho\tau_{81}$		8	1	8	$\frac{1}{8}$	0	8	8
$\rho\tau_{82}$		8	35	24	$\frac{1}{24}$	0	840	840
$\rho\tau_{91}$		9	1	9	$\frac{1}{9}$	$\frac{1}{2304}$	9	9
$\rho\tau_{92}$		9	56	45	$\frac{1}{45}$	$\frac{1}{11520}$	2520	2520
$\rho\tau_{10}$		10	1	10	$\frac{1}{10}$	0	10	10

Table 2: Non-superfluous rooted trees and related properties required to obtain order $p \leq 10$ with stages $s \leq 5$.

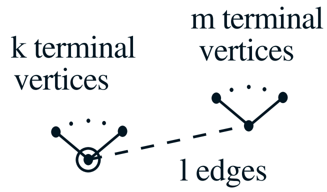


Figure 5: Analytic results exist for $\tilde{\delta}$ for trees of this type.

in addition to those depicted in table 2. However, for $s \leq 5$ the simplifying assumptions given in section 6 reduce all rooted trees of order $r \leq 10$ to those given in the table.

For larger s additional trees are necessary. Some of the properties are required in further related work and are provided here for completeness [32, 33]. The rooted trees we give are centroids, i.e., the root is shifted to a position minimizing the longest path from root to terminal vertex. There are at most two such positions for each tree [17], which are necessarily adjacent.

5 Determining $\tilde{\delta}$

The new density $\tilde{\delta}$ is more difficult to work with than the original δ since the latter is given by the relatively simple recursive scheme (2.1). In this section we present analytical results which yield the value of $\tilde{\delta}$ directly for certain types of trees. We consider a class of trees characterised by a triplet of natural numbers and identify a generating function[2] of three indeterminates such that its Taylor expansion provides the densities. In principle such generating functions can be constructed for any class of trees which is characterised by a finite number of natural numbers, although it is not clear how to find the functions in a simple and systematic way. Here we consider trees consisting of a slender part of l edges with bushy parts of k and m terminal vertices (leaves) at each end.

Definition 5.1 (Bushy and slender trees) *A* bushy tree *is a rooted tree where every child of the root is terminal. A* slender tree *is a rooted tree where each vertex has at most one child.*

In the usual terminology bushy trees are those of *height* 2, whereas the slender trees or those of *width* 1. We regard the trivial tree with $r(\rho\tau) = 1$ as both bushy and slender. Consider rooted trees of the type given in figure 5 and denote these by $\rho\tau^{(k,l,m)}$. With the notation $\tilde{\delta}^{(k,l,m)} = \tilde{\delta}(\rho\tau^{(k,l,m)})$ one obtains the following result:

Theorem 5.2 ($\tilde{\delta}$ for an important class of trees) *The new densities of the trees $\rho\tau^{(k,l,m)}$ are given by a multivariate Taylor expansion of the generating function*

[2]The expression *generating function* is used here in a generic sense and bears no relation to the concept of generating functions for symplectic transformations.

$$(5.1) \qquad g(t_1,t_2,t_3) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty} \frac{\tilde{\delta}^{(k,l,m)}}{k!\,m!}\, t_1^k t_2^l t_3^m =$$

$$\frac{t_1\ \sinh(\frac{t_1+t_2+t_3}{2})}{(t_1+t_2)\,(t_1+t_3)\cosh(\frac{t_2}{2})} + \frac{t_2\ \sinh(\frac{t_1+t_2-t_3}{2})}{(t_2+t_1)\,(t_2-t_3)\cosh(\frac{t_2}{2})} + \frac{t_3\ \sinh(\frac{t_1-t_2+t_3}{2})}{(t_3+t_1)\,(t_3-t_2)\cosh(\frac{t_2}{2})}\ .$$

Specialisations include

$$\begin{aligned}
&a)\quad g(0,t_2,t_3) &=& \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\frac{\tilde{\delta}^{(0,l,m)}}{m!}\,t_2^l t_3^m &=& \frac{2}{(t_2-t_3)}\frac{\sinh(\frac{t_2-t_3}{2})}{\cosh(\frac{t_2}{2})}\ ,\\
&b)\quad g(0,t_2,0) &=& \sum_{l=0}^{\infty}\tilde{\delta}^{(0,l,0)}\,t_2^l &=& \frac{2}{t_2}\tanh(\frac{t_2}{2})\ ,\\
&c)\quad g(0,0,t_3) &=& \sum_{m=0}^{\infty}\frac{\tilde{\delta}^{(0,0,m)}}{m!}\,t_3^m &=& \frac{2}{t_3}\sinh(\frac{t_3}{2})\ .
\end{aligned}$$

A sketch of the proof is given in the appendix. The values of $\tilde{\delta}$ provided by this theorem coupled with lemma 4.6 are sufficient to cover all the trees in table 2. The property

$$(5.2) \qquad g(t_1,t_2,t_3) = g(t_3,-t_2,t_1)\ , \quad \text{or equivalently,} \quad \tilde{\delta}^{(k,l,m)} = (-1)^l\ \tilde{\delta}^{(m,l,k)}$$

of the generating function reflects the symmetry of the trees, which are invariant under the exchange $k \leftrightarrow m$ together with shifting of the root along the slender part.

Furthermore, corollary 4.11 may be verified for the trees of class $\rho\tau^{(k,l,m)}$, since

$$g(t_1,t_2,t_3) = g(-t_1,-t_2,-t_3)\ , \text{ or equivalently, } \quad \tilde{\delta}^{(k,l,m)} = (-1)^{k+l+m}\ \tilde{\delta}^{(k,l,m)}\ .$$

Hence, for even order $r = k+l+m+1$ the density must vanish.

Example 5.3 The weights for a bushy m-tree are found to be

$$\tilde{\delta}^{(0,0,m-1)} = \begin{cases} \dfrac{1}{m\,2^{m-1}} & \text{for } m \text{ odd,}\\[2ex] 0 & \text{for } m \text{ even.}\end{cases}$$

For the $(k+m)$-tree $\rho\tau^{(k-1,1,m-1)}$, consisting of two bushy trees with $k-1$ and $m-1$ terminal vertices joined by an edge, one obtains

$$\tilde{\delta}^{(k-1,1,m-1)} = \begin{cases} \dfrac{-1}{k\,(k+m)\,2^{k+m-1}} & \text{for } k \text{ odd and } m \text{ even,}\\[2ex] \dfrac{1}{m\,(k+m)\,2^{k+m-1}} & \text{for } k \text{ even and } m \text{ odd,}\\[2ex] 0 & \text{for } k+m \text{ even.}\end{cases}$$

6 Modifying Butcher's simplifying assumptions.

In the search for Runge-Kutta schemes of maximal order, Butcher derived relations among the order conditions valid for implicit (not necessarily symplectic) schemes [5]. These simplifying assumptions may be applied to further reduce the number of rooted trees in the symplectic case. An algorithm for this process is presented in [24]. We re-formulate the results according to the new framework and demonstrate how the reduction in parameters in the new order condition formalism elucidates the attainable order of such methods.

Consider

$$\tilde{\Psi}_i(p) \;=\; \sum_{j=1}^{s} \tilde{a}_{ij}\, \tilde{c}_j^{p-1} \;-\; \frac{\tilde{c}_i^p}{p} \;+\; \frac{f_p}{p\,2^p}$$

and

$$\begin{aligned}
\tilde{o}(m) \;&=\; \sum_{i=1}^{s} b_i\, \tilde{c}_i^{m-1} \;-\; \tilde{\delta}^{(0,0,m-1)} \,, \\
\tilde{o}(k,m) \;&=\; \sum_{i,j=1}^{s} b_i\, \tilde{c}_i^{k-1}\, \tilde{a}_{ij}\, \tilde{c}_j^{m-1} \;-\; \tilde{\delta}^{(k-1,1,m-1)} \,,
\end{aligned}$$

with $\tilde{a}_{ij} = s_{ij} b_j$,

$$\tilde{c}_i = \sum_{j=1}^{s} \tilde{a}_{ij} = \sum_{j=1}^{s} s_{ij}\, b_j \tag{6.1}$$

and

$$f_p \;=\; \begin{cases} 1 & \text{for even } p, \\ 0 & \text{for odd } p. \end{cases}$$

The densities $\tilde{\delta}^{(\ldots)}$ are given in example 5.3. These quantities represent the new order conditions for bushy m-trees (i.e. $\tilde{o}(m) = 0$) and pairs of bushy k and m-trees joined by an edge (i.e. $\tilde{o}(k,m) = 0$). One finds the identity

$$\sum_{i=1}^{s} b_i\, \tilde{c}_i^{k-1}\, \tilde{\Psi}_i(p) \;=\; \tilde{o}(k,p) - \frac{1}{p}\, \tilde{o}(k+p) + \frac{f_p}{p\,2^p}\, \tilde{o}(k)$$

for arbitrary values of k and p. Following Butcher's arguments one considers $k = 1, \ldots, s$ and assumes that all order conditions in the right hand side vanish. The resulting system of homogeneous linear equations for $\tilde{\Psi}_i(p)$ involve the matrix $B = \mathrm{diag}(b_1, \ldots, b_s)$ and the Vandermonde matrix $(\tilde{c}_i^{k-1})$. They are invertible for b_i non-zero and $c_1, \ldots, c_s$ distinct, where we note that

$$\tilde{c}_i = c_i - \frac{1}{2} \sum_{j=1}^{s} b_j \;=\; c_i - \frac{1}{2}$$

under the consistency condition. Hence, in this case the order conditions

$$\tilde{B}(s+p): \quad \tilde{o}(1) = 0 \;,\; \ldots \;,\; \tilde{o}(s+p) = 0$$

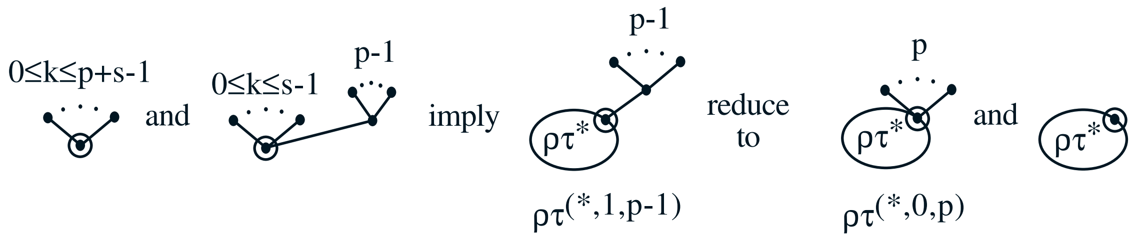


Figure 6: Illustration of the modified simplifying assumptions for an s-stage scheme.

for bushy trees together with the order conditions

$$\tilde{E}(s,p): \quad \tilde{o}(1,p)=0 \ , \ \ldots \ , \ \tilde{o}(s,p)=0$$

for trees of the type $\rho\tau^{(k-1,1,p-1)}$ imply $\tilde{\Psi}_i(p)=0$ for $i=1,\ldots,s$, i.e.

$$\sum_{j=1}^{s} \tilde{a}_{ij}\, \tilde{c}_j^{p-1} \;=\; \frac{\tilde{c}_i^p}{p} \;-\; \frac{f_p}{p\,2^p} \ . \tag{6.2}$$

The interpretation of this relation is as follows. The new weight $\tilde{\Phi}$ is constructed recursively from weight vectors $\tilde{\varphi}$ via (2.2) with a_{ij} replaced by $\tilde{a}_{ij}=a_{ij}b_j$. Then (6.2) states that the new weight vectors associated with the trees of section 5 are related in the following way:

$$\tilde{\varphi}_i(\rho\tau^{(0,1,p-1)}) \;=\; \frac{1}{p}\ \tilde{\varphi}_i(\rho\tau^{(0,0,p)}) - \frac{f_p}{p\,2^p} \ .$$

Consider the trees $\rho\tau^{(*,1,p-1)}$ and $\rho\tau^{(*,0,p)}$ consisting of an arbitrary tree $\rho\tau^*$ connected with a sub-tree $\tau^{(0,1,p-1)}$ and $\tau^{(0,0,p)}$ respectively (see figure 6). From the recursive construction one finds the following result for the total weights:

$$\tilde{A}(p): \quad \tilde{\Phi}(\rho\tau^{(*,1,p-1)}) \;=\; \frac{1}{p}\ \ \tilde{\Phi}(\rho\tau^{(*,0,p)}) - \frac{f_p}{p\,2^p}\ \ \tilde{\Phi}(\rho\tau^*) \ .$$

Using the principle of remark 4.5 this relation also holds for the density $\tilde{\delta}$. Hence, under the assumptions $\tilde{B}(s+p)$ and $\tilde{E}(s,p)$, the order conditions imposed by all trees of type $\rho\tau^{(*,1,p-1)}$ can be replaced by the order conditions associated with $\rho\tau^{(*,0,p)}$ and $\rho\tau^*$. This simplifying process, considered in conjunction with the invariance of the order conditions under root-shifting, gives rise to a considerable reduction in the number of trees.

Example 6.1 Assume that the order conditions of theorem 4.4 hold for trees up to order $p-1$, hence we have $\tilde{B}(p-1)$ and $\tilde{E}(s,p-s-1)$ and we may use the reduction $\tilde{A}(p-s-1)$. On adopting the necessary trees for $\tilde{B}(p)$ and $\tilde{E}(s,p-s)$ at pth order, we may use the $\tilde{A}(p-s)$ type reduction.

	Order										*Free*
	1	2	3	4	5	6	7	8	9	10	*Parameters*
$s=1$	1	1									1
$s=2$	1	1	2	3							3
$s=3$	1	1	2	3	5	6					6
$s=4$	1	1	2	3	6	$7^{\dagger}$	9	10			10
$s=5$	1	1	2	3	6	10	$10^{\ddagger}$	12	14	15	15

†The tree $\rho\tau_{53}$ from table 2 expires when order ≥ 6.
‡The trees $\rho\tau_{53}$, $\rho\tau_{63}$ and $\rho\tau_{64}$ from table 2 expire when order ≥ 7.

Table 3: Number of order conditions required for pth order, s-stage symplectic Runge-Kutta schemes.

Example 6.2 For a 4-stage 5th order symplectic Runge-Kutta method we obtain conditions for order 6 by adopting $\rho\tau_{61}$ and $\rho\tau_{62}$ from table 2. The tree $\rho\tau_{53}$ (necessary for a 4-stage 5th order method) reduces to $\rho\tau_{52}$ since $\tilde{A}(2)$ holds and the symplecticity condition enables shifting of the root. The tree $\rho\tau_{53}$ is said to *expire*. 7th order may be obtained by adopting $\rho\tau_{71}$ and $\rho\tau_{72}$ in addition to lower order trees, and 8th order may be obtained by further adopting $\rho\tau_{81}$.

Example 6.3 For a 5 stage 7th order symplectic Runge-Kutta method adopt the trees necessary for 8th order ($\rho\tau_{81}$ and $\rho\tau_{82}$ in table 2). Then $\tilde{A}(3)$ holds and $\rho\tau_{53}$, $\rho\tau_{63}$ and $\rho\tau_{64}$ expire (the order conditions are redundant since they are implied by imposing the other trees).

Thus table 2 gives sufficient rooted trees to derive all symplectic methods of order $p \leq 10$ with stages $s \leq 5$. To put this reduction into perspective, compare and contrast this with the number of order conditions in table 1. Illustration of the attainable order of symplectic Runge-Kutta schemes for stages $s \leq 5$ is given in table 3. The number of free parameters involved in the new formulation gives a clear indication of the attainable order $2s$ (the Gauss-Legendre schemes).

7 Application.

For a systematic application of the new framework developed here and a categorisation of symplectic Runge-Kutta methods we refer to [32]. In order to show that the parametrised Butcher-matrix yields a more natural description of symplectic methods, we provide the following example as illustration.

When implementing implicit Runge-Kutta schemes the majority of time is spent solving the nonlinear equations (1.2) for the intermediate stages $\mathbf{k}_i$ at each time step. Burrage [3] showed that considerable simplification of the equations occurs when the spectrum of A is real. Iserles [14] derived 3-stage 4th order symplectic Runge-Kutta methods with real spectra. Iserles and Nørsett provide bounds on the attainable order of such methods using order stars [15]. Here we provide a simple explicit proof of maximal order for a 2-stage method with real spectrum.

Theorem 7.1 *No 2-stage symplectic Runge-Kutta method of order $p \geq 3$ with real spectrum exists.*

Proof Assume the method is consistent, i.e. $b_1 + b_2 = 1$, then it is of second order since the only 2-tree is superfluous. The (single) additional condition for order 3 is

$$\sum_{i=1}^{2} b_i \Big(\sum_{j=1}^{2} s_{ij} b_j \Big)^2 = \frac{1}{12} \, .$$

With $s_{11} = s_{22} = 0$, $s_{12} = -s_{21} := x$ this order condition reduces to $x^2 b_1 b_2 = 1/12$. However, by virtue of consistency, the condition for real eigenvalues of the parametrised matrix (1.9) is given by $x^2 b_1 b_2 \leq 1/16$. $\square$

8 Conclusions and wider implications.

It has been demonstrated that the equations of Butcher's classical theory of order may be combined into an equivalent set of new order conditions which are more appropriate for symplectic and self-adjoint methods. The main advantages of the new formulation are as follows. The symplecticity condition can be implemented easily by assuming skew-symmetry of the parameter matrix S. This reduces the original number of free parameters by a factor of 2 and simultaneously reduces the number of order conditions, which are invariant under root-shifting. In contrast to the results in [28], for a given tree, this property is independent of lower order trees. Further, self-adjoint methods are given by a simple symmetry of the parameter matrix. In this case all order conditions associated with even order trees become redundant. Symplecticity and self-adjointness can be readily combined and arise as a consequence of persymmetry of the parameter matrix. This reduces the original number of free parameters by a factor of approximately 4. The order conditions reduce to those generated by odd non-rooted trees.

The condition $M = 0$ (see (1.8)) is important not just for symplecticity but also for many other desirable attributes of Runge-Kutta methods. For example it guarantees the correct reconstruction of Hopf bifurcations [16] and the conservation of orthogonality in unitary flows [7]. The condition $M \geq 0$ (often in tandem with $\mathbf{b} \geq 0$) arises in many aspects of the numerical analysis of ordinary differential equations. It has important consequences for algebraic stability and B-stability [34]. It is interesting to ask how these issues translate into the new formalism and what insight, if any, can be gained. Such questions are the focus of [32].

A computer algebra package for the derivation of order conditions and related properties for Runge-Kutta methods using Butcher's formalism is outlined in [31].

A significantly enhanced version of this package, `Butcher.m`, is available as part of release 3.0 of Mathematica onwards. We have also developed a package, built upon the functionality of `Butcher.m`, for the symplectic case. This has been accomplished by automating the formalism outlined here and is of particular relevance for several reasons. Derivation of $\tilde{\delta}$ is a much more involved computational process than for the standard density γ. Furthermore it is useful to automate the role of the stage-dependent symplifying assumptions. It is also useful to be able to fix the type of method under investigation (diagonally implicit, self-adjoint etc.) and to generate appropriate order conditions for methods within this class. The symplectic order condition package `SymplecticRungeKutta.m` can be obtained from either of the authors.

Acknowledgments

The authors gratefully acknowledge many useful discussions with Dr W. Schief. In particular, he pointed out the possibility of investigating the properties of the new density $\tilde{\delta}$ via the use of generating functions.

References

[1] P. Albrecht, *The Runge-Kutta Theory in a Nutshell*, SIAM J. Numer. Anal., 33 (1996) (to appear).

[2] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed., Springer, New York, 1989.

[3] K. Burrage, *A special family of Runge-Kutta methods for solving stiff differential equations*, BIT, 18 (1978), pp. 22–41.

[4] J. C. Butcher, *Implicit Runge-Kutta processes*, Math. Comp., 18 (1964), pp. 50–64.

[5] ——, *The Numerical Analysis of Ordinary Differential Equations*, John Wiley and Sons, Chichester, 1987.

[6] K. Dekker and J. G. Verwer, *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*, North Holland, Amsterdam, 1984.

[7] L. Dieci, B. Russell and E. van der Vleck, *Unitary integrators and Applications to Continuous Orthonormalization Techniques*, SIAM J. Numer. Anal., 31 (1994), pp. 261–281.

[8] P. Görtz and R. Scherer, *Reducibility and Characterization of Symplectic Runge-Kutta Methods*, ETNA, 2, (1994), pp. 194–204.

[9] H. Goldstein, *Classical Mechanics*, Second edition, Addison-Wesley, London, 1980.

[10] E. Hairer, S. P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems*, 2nd Edition, Springer, Berlin, 1993.

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

[12] F. Harary, *Graph Theory*, Addison-Wesley, London, 1969.

[13] M. E. Hosea, *A new recurrence for computing Runge-Kutta truncation error coefficients*, SIAM J. Numer. Anal., 32 (1995), pp. 1989–2001.

[14] A. Iserles, *Efficient Runge–Kutta methods for Hamiltonian systems*, Bull. Greek Math. Soc., 32 (1991), pp. 3–20.

[15] A. Iserles and S. P. Nørsett, *Order Stars*, Chapman and Hall, London, 1991.

[16] A. Iserles, A. T. Peplow and A. M. Stuart, *A unified approach to spurious solutions introduced by time descretisation. Part I: Basic theory*, SIAM J. Numer. Anal., 28 (1991), pp. 1723–1751.

[17] D. E. Knuth, *Fundamental Algorithms : The Art of Computer Programming*, Second edition, Addison-Wesley, London, 1973.

[18] F. Lasagni, *Canonical Runge–Kutta methods*, ZAMP, 39 (1988), pp. 952–953.

[19] R. S. MacKay, *Some aspects of the dynamics and numerics of Hamiltonian systems*, in Proceedings of the Dynamics of Numerics and Numerics of Dynamics Conference, D.S. Broomhead and A. Iserles, eds., Oxford University Press, Oxford, 1993.

[20] H. Munthe-Kaas, *Lie-Butcher theory for Runge-Kutta methods*, BIT, 35 (1995) pp. 572–587.

[21] S. P. Nørsett, *Collocation and perturbed collocation methods*, in Numerical Analysis, G.A. Watson, ed., Lecture Notes in Mathematics, 773, Springer, Berlin, 1980.

[22] S. P. Nørsett and G. Wanner, *Perturbed collocation*, Num. Math., 38 (1981), pp. 193–208.

[23] G. Ramaswami, *Perturbed collocation and symplectic RKN methods*, Adv. in Comp. Math. **3** 1 (1995), pp. 23–40.

[24] S. Saito, H. Sugiura and T. Mitsui, *Butcher's simplifying assumption for symplectic integrators*, BIT, 32 (1992), pp. 345–349.

[25] ——, *Family of implicit Runge-Kutta formulae*, BIT, 32 (1992), pp. 539–543.

[26] J. M. Sanz-Serna, *Runge-Kutta schemes for Hamiltonian systems*, BIT, 28 (1988), pp. 877–883.

[27] ——, *Symplectic integrators for Hamiltonian problems: an overview*, Acta Numerica, 1 (1992), pp. 243–286.

[28] J. M. Sanz-Serna and L. Abia, *Order conditions for canonical Runge-Kutta schemes*, SIAM J. Numer. Anal., 28 (1991), pp. 1081–1096.

[29] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, Chapman and Hall, London, 1994.

[30] M. Sofroniou, *Symbolic and Numerical Methods for Hamiltonian Systems*, Ph.D. thesis, Loughborough University of Technology (1994).

[31] ——, *Symbolic Derivation of Runge-Kutta Methods*, J. Symb. Comp., **18** 3, (1994) pp. 265–296.

[32] M. Sofroniou and W. Oevel, *Symplectic Runge-Kutta Schemes II: Classification*, (in preparation).

[33] ——, *Symplectic Runge-Kutta Schemes III: Canonical Elementary Differentials*, Mathematical Sciences Report A 204, Loughborough University of Technology, December 1993.

[34] A. M. Stuart, (1994). Numerical analysis of dynamical systems. Acta Numerica, **3**, pp. 467–572.

[35] Y. B. Suris, *Canonical transformations generated by methods of Runge–Kutta type for the numerical integration of the system* $\mathbf{x}'' = -\partial U/\partial \mathbf{x}$, Zh. Vychisl. Mat i Mat. Fiz., 29 (1989), pp. 202–211. (In Russian.); translated to English in U.S.S.R Comp. Maths. Math. Phy., 29 (1989), pp. 149–156.

Appendix

Proof of theorem 5.2

We derive recursive relations between the densities for different values of the parameters (k,l,m) which then lead to the identification of the generating function. We first establish b) (slender trees) and then use these results to derive a) (slender trees with a bushy canopy). Based on these formulae one can identify the general generating function (5.1). Part c) (bushy trees) is then obtained from the limit $t_2, t_3 \to 0$.

We introduce

$$\Phi^{(k,l,m)} = \Phi(\rho\tau^{(k,l,m)}) \quad , \quad \tilde{\Phi}^{(k,l,m)} = \tilde{\Phi}(\rho\tau^{(k,l,m)}) \quad , \quad \delta^{(k,l,m)} = \delta(\rho\tau^{(k,l,m)})$$

and

$$\tilde{A} = A - \mathbf{e}\mathbf{b}^T/2 = SB \ .$$

We use the parameters c_i and $\tilde{c}_i$ defined in (1.3) and (6.1), powers of which form the components of the vectors

$$\mathbf{c}^{(j)} = (c_1^j, \ldots, c_s^j)^T \ , \quad \tilde{\mathbf{c}}^{(j)} = (\tilde{c}_1^j, \ldots, \tilde{c}_s^j)^T \ .$$

The dimension of the vector, s, is the number of stages of the Runge-Kutta method. During the course of computations we make use of the property

$$< \mathbf{a}, \mathbf{b}\mathbf{c}^T\mathbf{d} > \quad = \quad < \mathbf{a}, \mathbf{b} >< \mathbf{c}, \mathbf{d} >$$

of the scalar product.

Proof of b) One finds the following recursion relation for the new weight function:

$$\begin{aligned}
\tilde{\Phi}^{(0,l,0)} &= < \mathbf{b}, \tilde{A}^l\mathbf{e} > \quad = \quad < \mathbf{b}, \tilde{A}^{l-1}(A - \frac{1}{2}\mathbf{e}\mathbf{b}^T)\mathbf{e} > \\
&= < \mathbf{b}, \tilde{A}^{l-2}(A - \frac{1}{2}\mathbf{e}\mathbf{b}^T)A\mathbf{e} > -\frac{1}{2} < \mathbf{b}, \tilde{A}^{l-1}\mathbf{e} >< \mathbf{b}, \mathbf{e} > \\
&= \cdots \\
&= < \mathbf{b}, A^l\mathbf{e} > -\frac{1}{2} < \mathbf{b}, \mathbf{e} >< \mathbf{b}, A^{l-1}\mathbf{e} > - \ldots - \frac{1}{2} < \mathbf{b}, \tilde{A}^{l-1}\mathbf{e} >< \mathbf{b}, \mathbf{e} > \\
&= \Phi^{(0,l,0)} - \frac{1}{2}\sum_{j=0}^{l-1} \tilde{\Phi}^{(0,l-j-1,0)}\Phi^{(0,j,0)} \ . \qquad \text{(A.1)}
\end{aligned}$$

Using remark 4.5 we may replace the weight functions by the densities

$$\Phi^{(0,i,0)} = \delta^{(0,i,0)} = \frac{1}{(i+1)!} \quad \text{and} \quad \tilde{\Phi}^{(0,i,0)} = \tilde{\delta}^{(0,i,0)}.$$

Summing both sides of (A.1) one obtains

$$\sum_{l=0}^{\infty} \tilde{\delta}^{(0,l,0)}\, t_2^l = \sum_{l=0}^{\infty} \frac{t_2^l}{(l+1)!} - \frac{1}{2}\sum_{l=0}^{\infty}\sum_{j=0}^{l-1} \frac{\tilde{\delta}^{(0,l-j-1,0)}}{(j+1)!}\, t_2^l \ .$$

Next, transform the range of summation using $p = l - j - 1$:

$$\begin{aligned}
\sum_{l=0}^{\infty} \tilde{\delta}^{(0,l,0)}\, t_2^l &= \left(\frac{\exp(t_2) - 1}{t_2}\right) - \frac{1}{2}\sum_{j=0}^{\infty}\sum_{p=0}^{\infty} \frac{\tilde{\delta}^{(0,p,0)}}{(j+1)!}\, t_2^{p+j+1} \\
&= (\exp(t_2) - 1)\left(\frac{1}{t_2} - \frac{1}{2}\sum_{p=0}^{\infty} \tilde{\delta}^{(0,p,0)}\, t_2^p\right) .
\end{aligned}$$

Re-arranging we obtain:

$$\sum_{l=0}^{\infty} \tilde{\delta}^{(0,l,0)}\, t_2^l = \frac{2}{t_2}\frac{(\exp(t_2) - 1)}{(\exp(t_2) + 1)}$$

and the result follows.

Proof of a) Using similar arguments one derives the recursion relation:

$$\begin{aligned}
\tilde{\Phi}^{(0,l,m)} &= \sum_{i=1}^{s} b_i(\tilde{A}^l\tilde{\mathbf{c}}^{(m)})_i \quad = \quad <\mathbf{b}, \tilde{A}^{l-1}(A-\frac{1}{2}\mathbf{e}\mathbf{b}^T)\tilde{\mathbf{c}}^{(m)}> \\
&= <\mathbf{b}, \tilde{A}^{l-2}(A-\frac{1}{2}\mathbf{e}\mathbf{b}^T)A\tilde{\mathbf{c}}^{(m)}> -\frac{1}{2}<\mathbf{b}, \tilde{A}^{l-1}\mathbf{e}><\mathbf{b}, \tilde{\mathbf{c}}^{(m)}> \\
&= \cdots \\
&= <\mathbf{b}, A^l\tilde{\mathbf{c}}^{(m)}> -\frac{1}{2}\sum_{j=0}^{l-1}\tilde{\Phi}^{(0,l-j-1,0)}<\mathbf{b}, A^j\tilde{\mathbf{c}}^{(m)}> \quad . \qquad \text{(A.2)}
\end{aligned}$$

Substituting

$$<\mathbf{b}, A^j\tilde{\mathbf{c}}^{(m)}> \;=\; <\mathbf{b}, A^j\sum_{i=0}^{m}\binom{m}{i}\left(-\frac{1}{2}\right)^{m-i}\mathbf{c}^{(i)}> \;=\; \sum_{i=0}^{m}\binom{m}{i}\left(-\frac{1}{2}\right)^{m-i}\Phi^{(0,j,i)} \; .$$

into (A.2) gives:

$$\text{(A.3)}\quad \tilde{\Phi}^{(0,l,m)} = \sum_{i=0}^{m}\binom{m}{i}\left(-\frac{1}{2}\right)^{m-i}\Phi^{(0,l,i)} + \sum_{j=0}^{l-1}\sum_{i=0}^{m}\binom{m}{i}\left(-\frac{1}{2}\right)^{m-i+1}\tilde{\Phi}^{(0,l-j-1,0)}\Phi^{(0,j,i)} \; .$$

Using remark 4.5 one may replace

$$\Phi^{(0,i,j)} = \delta^{(0,i,j)} = \frac{j!}{(i+j+1)!} \quad \text{and} \quad \tilde{\Phi}^{(0,i,j)} = \tilde{\delta}^{(0,i,j)} \; .$$

Summing both sides of (A.3) and using the re-summations $i \to m-i$ and $p = l-j-1$ one identifies the generating function found in b) in the right hand side:

$$\begin{aligned}
\frac{1}{m!}\sum_{l=0}^{\infty}\tilde{\delta}^{(0,l,m)}\, t_2^l \quad &= \quad \sum_{l=0}^{\infty}\frac{t_2^l}{(m+l+1)!}\sum_{i=0}^{m}\binom{m+l+1}{i}\left(-\frac{1}{2}\right)^i \\
&\quad +\Big(\sum_{p=0}^{\infty}\tilde{\delta}^{(0,p,0)}\, t_2^p\Big)\;\sum_{j=0}^{\infty}\sum_{i=0}^{m}\binom{m+j+1}{i}\left(-\frac{1}{2}\right)^{i+1}\frac{t_2^{j+1}}{(m+j+1)!} \; .
\end{aligned}$$

Insertion of the result b), summation over m and various re-summations lead to the stated result.

Proof of (5.1) A recursion relation is obtained from

$$\begin{aligned}
\tilde{\Phi}^{(k,l,m)} &= \sum_{i=1}^{s} b_i\tilde{c}_i^{(k)}(\tilde{A}^l\tilde{\mathbf{c}}^{(m)})_i \quad = \quad <B\tilde{\mathbf{c}}^{(k)}, \tilde{A}^l\tilde{\mathbf{c}}^{(m)}> \\
&= <B\tilde{\mathbf{c}}^{(k)}, \tilde{A}^{l-1}(A-\frac{1}{2}\mathbf{e}\mathbf{b}^T)\tilde{\mathbf{c}}^{(m)}> \\
&= <B\tilde{\mathbf{c}}^{(k)}, \tilde{A}^{l-2}(A-\frac{1}{2}\mathbf{e}\mathbf{b}^T)A\tilde{\mathbf{c}}^{(m)}> -\frac{1}{2}<B\tilde{\mathbf{c}}^{(k)}, \tilde{A}^{l-1}\mathbf{c}^{(0)}><\mathbf{b}, \tilde{\mathbf{c}}^{(m)}> \\
&= \cdots \\
&= <B\tilde{\mathbf{c}}^{(k)}, A^l\tilde{\mathbf{c}}^{(m)}> -\frac{1}{2}\sum_{j=0}^{l-1}\tilde{\Phi}^{(k,l-j-1,0)}<\mathbf{b}, A^j\tilde{\mathbf{c}}^{(m)}> \quad .
\end{aligned}$$

Using

$$< \mathbf{b}, A^j \tilde{\mathbf{c}}^{(m)} > \;=\; \sum_{r=0}^{m} \binom{m}{r} \left(-\frac{1}{2}\right)^{m-r} < \mathbf{b}, A^j \mathbf{c}^{(r)} > \;=\; \sum_{r=0}^{m} \binom{m}{r} \left(-\frac{1}{2}\right)^{m-r} \Phi^{(0,j,r)}$$

and

$$\begin{aligned} < B\tilde{\mathbf{c}}^{(k)}, A^l \tilde{\mathbf{c}}^{(m)} > &= \sum_{r_1=0}^{k} \sum_{r_2=0}^{m} \binom{k}{r_1} \binom{m}{r_2} \left(-\frac{1}{2}\right)^{k+m-r_1-r_2} < \mathbf{b}\mathbf{c}^{(r_1)}, A^l \mathbf{c}^{(r_2)} > \\ &= \sum_{r_1=0}^{k} \sum_{r_2=0}^{m} \binom{k}{r_1} \binom{m}{r_2} \left(-\frac{1}{2}\right)^{k+m-r_1-r_2} \Phi^{(r_1,l,r_2)} \end{aligned}$$

one finds

$$\begin{aligned} \tilde{\Phi}^{(k,l,m)} &= \sum_{r_1=0}^{k} \sum_{r_2=0}^{m} \binom{k}{r_1} \binom{m}{r_2} \left(-\frac{1}{2}\right)^{k+m-r_1-r_2} \Phi^{(r_1,l,r_2)} \\ &\quad -\frac{1}{2} \sum_{j=0}^{l-1} \tilde{\Phi}^{(k,l-j-1,0)} \sum_{r=0}^{m} \binom{m}{r} \left(-\frac{1}{2}\right)^{m-r} \Phi^{(0,j,r)} \;. \end{aligned} \tag{A.4}$$

Replacing the weights by the densities

$$\Phi^{(h,i,j)} = \delta^{(h,i,j)} = \frac{j!}{(i+j)!\,(h+i+j+1)} \quad \text{and} \quad \tilde{\Phi}^{(h,i,j)} = \tilde{\delta}^{(h,i,j)} \;,$$

using $\tilde{\delta}^{(h,i,0)} = (-1)^i \tilde{\delta}^{(0,i,h)}$ (5.2) and summing over the free indices one can identify the generating function a) in the right hand side of (A.4). Various re-summations yield (5.1). Details of the computation are given in [30]. □