An Analysis of The Order of Runge-Kutta Methods That Use an Iterative Scheme To Compute Their Internal Stage Values*

K. R. Jackson†, A. Kværnø‡ and S. P. Nørsett§

Submitted to BIT
May 1994

Abstract

We employ B-series to analyze the order of Runge-Kutta (RK) methods that use simple iteration, modified Newton iteration or full Newton iteration to compute their internal stage values. Our assumptions about the initial guess for the internal stage values cover a wide range of practical schemes. Moreover, the analytical techniques developed in this paper can be applied in many other similar contexts.

Key words. Numerical Solution, Initial Value Problem, IVP, Ordinary Differential Equation, ODE, Runge-Kutta Method, Order, Internal Stage Values, Iterative Scheme, Newton's Method.

AMS(MOS) subject classification. 65L05.

---

* This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Information Technology Research Centre of Ontario, and the Norwegian Research Council for Science and Humanities.

† Computer Science Department, University of Toronto, Toronto, Ontario, Canada, M5S 1A4. E-mail: krj@cs.toronto.edu.

‡ Division of Mathematical Sciences, Norwegian Institute of Technology, N–7034 Trondheim–NTH, Norway. E-mail: anne@imf.unit.no.

§ Division of Mathematical Sciences, Norwegian Institute of Technology, N–7034 Trondheim–NTH, Norway. E-mail: norsett@imf.unit.no.

1 Introduction.

Consider the Initial Value Problem (IVP) for Ordinary Differential Equations (ODEs)

$$y'(x) = f(y(x)), \qquad y(x_0) = y_0, \tag{1.1}$$

where x, y, $f \in \mathbb{R}$. Although all our results apply to systems of nonautonomous IVPs, to simplify the notation, we restrict the discussion throughout this paper to scalar autonomous problems.

An s-stage Runge-Kutta (RK) formula for the numerical solution of (1.1) is given by

$$Y_{n,i} = y_n + h \sum_{j=1}^{s} a_{ij} f(Y_{n,j}), \qquad i = 1, \cdots, s, \tag{1.2}$$

$$y_{n+1} = y_n + h \sum_{j=1}^{s} b_j f(Y_{n,j}), \tag{1.3}$$

where $\{Y_{n,i}\}$ are the *internal stage values* and $y_n \approx y(x_n)$. Formulas (1.2) and (1.3) can be written more compactly as

$$Y_n = y_n \varepsilon_s + hAf(Y_n) \tag{1.4}$$

$$y_{n+1} = y_n + hb^T f(Y_n) \tag{1.5}$$

where, for a vector v, v^T is its transpose, $\varepsilon_s = (1, \ldots, 1)^T \in \mathbb{R}^s$ is the vector of s ones, $Y_n = (Y_{n,1}, \ldots, Y_{n,s})^T$ is the vector of internal stage values, $A = [a_{ij}]$ is the matrix of RK *coefficients*, and $b = (b_1, \ldots, b_s)^T$ is the vector of RK *weights*. We use f for both the function associated with the IVP (1.1) and also that function applied to the components of a vector, namely, $f(Y_n) = (f(Y_{n,1}), \ldots, f(Y_{n,s}))^T$. Although not mentioned above, we will soon need the RK *nodes* $c_i = \sum_{j=1}^{s} a_{ij}$, also defined by $c = A\varepsilon_s$ for $c = (c_1, \ldots, c_s)^T$. In addition to the order of the formula, we will need its *stage order*, defined to be the largest integer η for which the C *simplifying assumptions*

$$C(\eta): \qquad \sum_{j=1}^{s} a_{ij} c_j^{r-1} = \frac{1}{r} c_i^r, \qquad i = 1, \ldots, s, \qquad r = 1, \ldots, \eta, \tag{1.6}$$

hold. Note that the stage order $\eta \geq 1$, since, from the definition of the nodes c_i above, $\sum_{j=1}^{s} a_{ij} = c_i$ for $i = 1, \ldots, s$.

For the past several years, we have been investigating parallel RK methods. One of the first schemes we considered was the RK *predictor-corrector* (RK–PC) formula

$$Y_n^{(k+1)} = y_n \varepsilon_s + hAf(Y_n^{(k)}), \tag{1.7}$$

$$y_{n+1}^{(k+1)} = y_n + hb^T f(Y_n^{(k)}), \tag{1.8}$$

which corresponds to *simple iteration* applied to (1.4)–(1.5). If this iteration is performed for $k = 0, \ldots, K-1$, then in practice we take $y_{n+1} = y_{n+1}^{(K)}$ and proceed to the next step. Clearly, if our goal is simply to compute $y_{n+1}^{(K)}$, we need not compute $y_{n+1}^{(k)}$ for $k < K$ or $Y_n^{(k)}$

for $k = K$, but, for the theoretical analysis to follow, it is convenient to define these values by (1.7)–(1.8). Moreover, these extra values might be useful for other purposes, such as error estimation or computing $Y_{n+1}^{(0)}$.

RK–PC schemes have the advantage that the s function evaluations associated with $f(Y_n^{(k)})$ in (1.7) can be performed in parallel and then whichever of the components of $Y_n^{(k+1)}$ and $y_n^{(k+1)}$ that are needed can be computed simultaneously, making RK–PC methods attractive for implementation on parallel computers having a few CPUs and fast interprocessor communication, properties characteristic of shared memory multiprocessors. A more complete discussion of RK–PC schemes can be found in [7, 9, 10, 12, 13, 14, 16].

One of us has also recently used RK–PC schemes to solve systems of hyperbolic equations by the method of characteristics. Even for serial computation on this class of problems, RK–PC schemes have important advantages not shared by more conventional ODE methods, as explained in more detail in [1, 5].

In our investigation of RK–PC schemes, we considered initial guesses $Y_n^{(0)}$ for Y_n satisfying

$$Y_{n,i}^{(0)} = y(x_n + c_i h; x_n, y_n) + O(h^{q+1}), \tag{1.9}$$

where $y(x; x_n, y_n)$ is the solution to the *local* IVP

$$y' = f(y), \qquad y(x_n; x_n, y_n) = y_n. \tag{1.10}$$

Such initial guesses can be computed in a variety of ways. For example, one could use an interpolant based on passed values or a low-order RK method to compute each $Y_{n,i}^{(0)}$; both approaches can be implemented effectively in parallel. Using the Lipschitz continuity of f, it is easy to prove that, if

$$Y_n^{(0)} - Y_n = O(h^{q+1}) \tag{1.11}$$

then the local error satisfies

$$y_{n+1}^{(k)} - y(x_{n+1}; x_n, y_n) = O(h^{\min(p,q+k)+1}), \tag{1.12}$$

where p is the order of the underlying RK formula. However, although it is easy to compute initial guesses satisfying (1.9) for q large, this is difficult for (1.11), unless the stage order of the underlying RK formula is at least q. Fortunately, we discovered the rather remarkable result that, if $Y_n^{(0)}$ satisfies (1.9), then the local error satisfies (1.12) even though $Y_n^{(0)}$ might be a poor initial approximation to Y_n in the sense that the largest μ for which

$$Y_n^{(0)} - Y_n = O(h^{\mu+1})$$

is much smaller than q.

We developed a theoretical explanation for this order behaviour based on *B-series* (sometimes called *Butcher series*). A more tutorial and less rigorous overview of our approach is presented in [11]. Burrage [2] adapts our analysis to determine the error coefficients for $y_{n+1}^{(k)} - y(x_{n+1}; x_n, y_n)$ in his study of RK–PC schemes and Kværnø [15] extends our approach to differential algebraic equations (DAEs).

In this paper, we consider the more general class of *iterated RK methods*

$$\begin{aligned}
\left(I_s - hAJ_n^{(k)}\right)\left(Y_n^{(k+1)} - Y_n^{(k)}\right) &= y_n\varepsilon_s + hAf(Y_n^{(k)}) - Y_n^{(k)}, && (1.13)\\
y_{n+1}^{(k+1)} &= y_n + hb^T f(Y_n^{(k)}), && (1.14)
\end{aligned}$$

where $k \geq 0$, I_s is the $s \times s$ identity matrix and $J_n^{(k)}$ is some approximation to the Jacobian associated with (1.4). Iteration (1.13) can be viewed as an application of a variant of Newton's method to solve (1.4), whereas (1.14) is identical to (1.8) used in simple iteration.

We could replace (1.14) by

$$y_{n+1}^{(k+1)} = y_n + hb^T f(Y_n^{(k)}) + hb^T J_n^{(k)}\left(Y_n^{(k+1)} - Y_n^{(k)}\right). \tag{1.15}$$

Iteration (1.13) and (1.15) together correspond to applying the same variant of Newton's method to solve the larger system (1.4)–(1.5). However, computing $y_{n+1}^{(K)}$ by (1.15) requires one more evaluation of (1.13) than does computing $y_{n+1}^{(K)}$ by (1.14), since (1.15) uses both $Y_n^{(K)}$ and $Y_n^{(K-1)}$, while (1.14) requires $Y_n^{(K-1)}$ only. Therefore, although the $y_{n+1}^{(K)}$ computed by (1.15) might be of higher-order than the one computed by (1.14), it also requires more work. In fact, computing $y_{n+1}^{(K)}$ by (1.15) requires about the same amount of work as computing $y_{n+1}^{(K+1)}$ by (1.14), and the latter is normally of at least as high an order. Nevertheless, (1.15) might be a useful alternative to (1.14) in some contexts.

If the RK coefficient matrix A is nonsingular, another alternative to (1.14) is

$$y_{n+1}^{(k+1)} = y_n + b^T A^{-1}(Y^{(k)} - y_n\varepsilon_s) \tag{1.16}$$

which follows from (1.14) and the observation that (1.4) can be rewritten as

$$hf(Y_n) = A^{-1}\left(Y_n - y_n\varepsilon_s\right).$$

The only difference between (1.16) and (1.14) is that in (1.16) the additional fixed point iteration employed in (1.14) is not performed, usually reducing the order of (1.16) by 1 compared to (1.14) when the order of $y_{n+1}^{(k+1)}$ is less than the order of the underlying RK formula (1.4)–(1.5).

Largely because of space constraints, we do not explicitly analyze the iteration (1.13) together with (1.15) or (1.16), although such an analysis follows easily from the results in this paper. Instead, we analyze in detail below the iteration (1.13)–(1.14) with three particular choices of $J_n^{(k)}$:

Case 1: $J_n^{(k)} = 0$ — *simple iteration*, making (1.13)–(1.14) equivalent to (1.7)–(1.8),

Case 2: $J_n^{(k)} = f'(y_n)I_s$ — a *modified Newton iteration*, and

Case 3: $J_n^{(k)} = \mathrm{diag}(f'(Y_{n,1}^{(k)}), \ldots, f'(Y_{n,s}^{(k)}))$ — the *full Newton iteration.*

The general scheme (1.13)–(1.14) shares with (1.7)–(1.8) the advantage that all function evaluations associated with $f(Y_n^{(k)})$ can be evaluated simultaneously. Thus, as discussed in

[12], the effective implementation of parallel RK methods in Cases 2 or 3 depends on being able to solve the linear system (1.13) effectively in parallel.

Iterative schemes of the form (1.13)–(1.14) are also used to approximate the solution of (1.4)–(1.5) in standard sequential codes based on implicit RK formulas. Both STRIDE [3] and SIMPLE [17] use simple iteration to solve (1.4)–(1.5) in the nonstiff intervals of the integration and a variant of modified Newton in the stiff intervals. Thus gaining a better understanding of the order of $y_{n+1}^{(k)}$ will give us more insight into standard sequential codes as well as novel parallel RK methods.

In the next section, we introduce definitions and preliminary results that allow us to use B-series in §3 to determine the order of $y_{n+1}^{(k)}$ and the associated error coefficients based on

- the order, and possibly the stage order, of the underlying formula (1.4)–(1.5),
- the initial guess $Y_n^{(0)}$,
- the choice of $J_n^{(k)}$, and
- the iteration count k.

The following central result is proved in §3.

Theorem 1.1 *Let the underlying RK formula be of order $p \geq 1$. If the initial guess $Y_n^{(0)}$ satisfies (1.9) with $q \geq -1$, then*

$$y_{n+1}^{(k)} = y(x_{n+1}; x_n, y_n) + O(h^{\min(p,q_k)+1}) \tag{1.17}$$

for $k \geq 1$, where q_k satisfies

Case 1: $q_k = q + k$ *if* $J_n^{(k)} = 0$,

Case 2: $q_k = q + 2k - 1$ *for* $q \geq 0$ *and* $q_k = \max(0, 2k-3)$ *for* $q = -1$ *if* $J_n^{(k)} = f'(y_n) I_s$, *and*

Case 3: $q_k = 2^{k-1}(r+2) + (q-r) - 1$ *for* $r = \min(q, \eta)$ *if* $J_n^{(k)} = diag(f'(Y_{n,1}^{(k)}), \ldots, f'(Y_{n,s}^{(k)}))$ *and the underlying RK formula satisfies $C(\eta)$.*

We always benefit from a high-order predictor, even if the stage order of the underlying formula is low, but, in the full Newton case, the advantage of q large might be restricted by the stage order η. In §3.3, we prove that, if the initial guess $Y_n^{(0)}$ satisfies a slightly stronger assumption than (1.9), then r can be replaced by q in Case 3 of Theorem 1.1, eliminating the restriction noted above.

The results established in §3 are valid for more general predictors than (1.9) and can be used in a variety of other ways, one of which is to determine the error coefficients associated with $y_{n+1}^{(k)}$. We do not, though, pursue the latter topic in detail in this paper.

In addition, the theory developed in §3 can be used to explain some possibly surprising order behaviour, an instance of which is the *order increase* phenomenon described in Example 2 of [11]. Another such peculiarity is that the order of $y_{n+1}^{(k)}$ might decrease as k increases, as shown in Example 4.5 of §4. Other numerical results illustrating our theory are presented in §4. We end with a few concluding remarks in §5.

The techniques developed in this paper can be applied in many other similar contexts. For example, as noted above, they can be used to compute the error coefficients of a RK–PC method to any order desired. Also, our techniques can be used to analyze similar schemes, such as the *PDIRK* methods of van der Houwen, Sommeijer and Couzy [9]. Moreover, our approach can be extended to allow for perturbed Jacobians in Cases 2 and 3 of Theorem 1.1. This extension allows us, for example, to apply our techniques to analyze practical methods that approximate $J_n^{(k)}$ by differences or use an approximate or exact Jacobian computed on an earlier iteration or a previous step. We plan to investigate these and other topics in a forthcoming paper.

2 Some Notation, Definitions and Preliminary Results.

As noted in §1, we use B-series to prove Theorem 1.1. However, before giving the proof in §3, we first introduce some necessary notation and provide a few definitions and preliminary results in this section. Some of these results are new while others have appeared before. See [4] or [8] for a more comprehensive introduction to B-series.

To begin, we introduce two types of trees that are fundamental to the theory of B-series. A *rooted tree* is a tree with one distinguished node, called the root. A *monotonically labeled rooted tree* (sometimes called a *labeled rooted tree* or simply a *labeled tree*) is a rooted tree whose nodes are labeled by an ordered set such that the label on each node is less than the label on any of its sons. We use T to denote the set of all monotonically labeled rooted trees.

We use the *bracket* notation of Butcher to designate particular trees: $\emptyset$ is the empty tree with no nodes, $\tau = [\emptyset]$ is the tree with one node, and, if $t_1, \ldots, t_m$ are rooted trees, then $t = [t_1, \ldots, t_m]$ is the rooted tree formed by joining the root of each tree t_i to a new common root for t. We use $[_k$ and $]_k$ to represent k left or right brackets, respectively, and t^k to denote k copies of the tree t, with the convention that $t^0 = \emptyset$. Some trees are illustrated in Example 2.1 and Figure 2.1.

The *order* $\rho(t)$ of a tree t is the number of nodes in t. Thus, $\rho(\emptyset) = 0$, $\rho(\tau) = 1$ and $\rho([t_1, \ldots, t_m]) = 1 + \sum_{i=1}^{m} \rho(t_i)$.

Also fundamental to the theory of B-series are the *elementary differentials* $F(t)(y)$ of $f(y)$ defined recursively by

$$\begin{aligned} F(\emptyset)(y) &= y, \\ F(\tau)(y) &= f(y), \\ F([t_1, \ldots, t_m])(y) &= f^{(m)}(y) F(t_1)(y), \ldots, F(t_m)(y) \end{aligned} \tag{2.1}$$

where $\rho(t_i) \geq 1$ for $i = 1, \ldots, m$ and $f^{(m)}(y) = d^m f(y)/dy^m$.

Let

$$G = \{a : T \to \mathbb{R}^l\}$$

where l is a positive integer, in our case usually 1, s or $s+1$.

Definition 2.1 *Let $a \in G$. The series*

$$B(a, y) = \sum_{t \in T} a(t) F(t)(y) \frac{h^{\rho(t)}}{\rho(t)!}$$

is called a B-series (*or sometimes a* Butcher series).

The importance of B-series stems from the well-known result

Theorem 2.1

$$B(a, y_n) = \sum_{t \in T} a(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} = y(x_n + \alpha h; x_n, y_n) + O(h^{p+1})$$

if and only if

$$a(t) = \alpha^{\rho(t)} \quad \forall\, t \in T \ s.t.\ \rho(t) \leq p.$$

If $u = (u_1, \ldots, u_l)^T$ and $v = (v_1, \ldots, v_l)^T$ are two vectors in $\mathbb{R}^l$, let uv be the element-by-element product of u and v: $uv = (u_1v_1, \ldots, u_lv_l)^T \in \mathbb{R}^l$. Similarly, if $w_i \in \mathbb{R}^l$ for $i = 1, \ldots, m$, let $\prod_{i=1}^m w_i = w_1 w_2 \cdots w_m \in \mathbb{R}^l$. For $m < 1$, define $\prod_{i=1}^m w_i = \varepsilon_l$. Moreover, if $w_i = w$ for $i = 1, \ldots, m$, then denote $\prod_{i=1}^m w_i$ by w^m with $w^0 = \varepsilon_l$. Note uv commutes, that is, $uv = vu$, so the order of the vectors in $\prod_{i=1}^m w_i$ is immaterial.

We can now state the well-known key result

Theorem 2.2 *If $a \in G$ and $a(\emptyset) = \varepsilon_l$, then*

$$hf(B(a, y)) = B(a', y)$$

where

$$\begin{aligned} a'(\emptyset) &= 0 \\ a'(\tau) &= \varepsilon_l \\ a'(t = [t_1, \ldots, t_m]) &= \rho(t) \prod_{i=1}^m a(t_i) \quad \textit{for } \rho(t) \geq 2. \end{aligned}$$

See [8, p. 246] for a proof.

Nørsett and Wolfbrandt prove the following result in [18, Theorem 2.2].

Theorem 2.3 *Let $B(a, y)$ and $B(b, y)$ be two B-series with a, $b \in G$, $a(\emptyset) = \varepsilon_l$ and $b(\emptyset) = 0$. Then*

$$hf'(B(a, y))\, B(b, y) = B(a \circ b, y)$$

where

$$\begin{aligned} (a \circ b)(\emptyset) &= 0 \\ (a \circ b)(\tau) &= 0 \\ (a \circ b)(t = [t_1, \ldots, t_m]) &= \rho(t) \sum_{i=1}^m \left(\prod_{\substack{j=1 \\ j \neq i}}^m a(t_j) \right) b(t_i) \quad \textit{for } \rho(t) \geq 2, \end{aligned} \tag{2.2}$$

where the product on the right side of (2.2) is ε_l if $m = 1$.

It is sometimes convenient to write (1.4)–(1.5) more compactly as

$$\tilde{Y}_n = y_n \varepsilon_{s+1} + h\tilde{A} f(\tilde{Y}_n) \tag{2.3}$$

where

$$\tilde{Y}_n = (Y_{n,1}, \ldots, Y_{n,s}, y_{n+1})^T \qquad \text{and} \qquad \tilde{A} = \begin{bmatrix} A & 0 \\ b^T & 0 \end{bmatrix}.$$

Thus, $y_{n+1} = \tilde{Y}_{n,s+1}$. The extended c-vector is $\tilde{c} = \tilde{A}\varepsilon_{s+1}$. Note $\tilde{c}_i = c_i = \sum_{j=1}^s a_{ij}$ for $i = 1, \ldots, s$ and $\tilde{c}_{s+1} = \sum_{i=1}^s b_i$. Since we assume throughout this paper that the RK formula is consistent (i.e., of order at least 1), $\tilde{c}_{s+1} = 1$. The size of the system is the only thing that distinguishes (1.4) from the *extended formula* (2.3). Thus, results given for one formulation are usually valid for the other as well. Throughout this paper, we reserve the $\tilde{}$ to mark values associated with the extended formula (2.3).

Before stating the next result concerning the B-series for the RK formula (2.3), we must define the multiplication of a product of vectors by a matrix. If $A \in \mathbb{R}^{l \times l}$ and $w_i \in \mathbb{R}^l$ for $i = 1, \ldots, m$, let

$$A \prod_{i=1}^{m} w_i = A \left(\prod_{i=1}^{m} w_i \right).$$

That is, form the element-by-element product of the vectors before multiplying by A. Note that in general $A(uv) \neq (Au)v$. As an example of the use of this notation, we rewrite the C simplifying assumptions (1.6) in the more compact form

$$C(\eta): \qquad Ac^{r-1} = \frac{1}{r} c^r \qquad \text{for } r = 1, \ldots, \eta. \tag{2.4}$$

As noted in §1, the stage order $\eta \geq 1$, since $Ac^0 = A\varepsilon_s = c$ from the convention that $w^0 = \varepsilon_s$ for any vector w and the definition of c, namely $c = A\varepsilon_s$.

Using this matrix-vector notation, we can now state the well-known fundamental result

Theorem 2.4 *Let $\tilde{Y}_n = B(\tilde{\phi}, y_n)$ be the B-series for the RK formula (2.3). Then $\tilde{\phi}(t)$ satisfies*

$$\begin{aligned} \tilde{\phi}(\emptyset) &= \varepsilon_{s+1} \\ \tilde{\phi}(t = [t_1, \ldots, t_m]) &= \tilde{A}\tilde{\phi}'(t) = \rho(t)\tilde{A} \prod_{i=1}^{m} \tilde{\phi}(t_i) \quad \text{for } \rho(t) \geq 1. \end{aligned}$$

Moreover, the RK formula is of order p if and only if p is the largest integer such that

$$\tilde{\phi}_{s+1}(t) = 1 \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq p.$$

Note that it immediately follows from Theorem 2.4 that $\tilde{\phi}(\tau) = \tilde{A}\tilde{\phi}(\emptyset) = \tilde{A}\varepsilon_{s+1} = \tilde{c}$.

We use the last theorem together with (1.6) (or equivalently (2.4)) to prove the following well-known result, which is used in §3.3.

Theorem 2.5 *The stage order of a RK formula is the largest integer η such that*

$$\phi(t) = c^{\rho(t)} \qquad \forall\, t \in T \text{ s.t. } \rho(t) \leq \eta. \tag{2.5}$$

Proof: Recall that the stage order η of a RK formula is defined to be the largest integer η such that (2.4) holds. We need to show that η is also the largest integer such that (2.5) holds. We establish this result by showing that (2.4) holds for a given value of $\eta \geq 1$ if and only if (2.5) holds for the same value of η.

Suppose (2.4) holds for some $\eta \geq 1$. We show by induction on $\rho(t)$ that (2.5) holds too for the same value of η. If $\rho(t) = 0$, then $t = \emptyset$ and $\phi(\emptyset) = \varepsilon_s = c^0$ from Theorem 2.4 and the convention noted earlier that $c^0 = \varepsilon_s$. Now assume that $\phi(t) = c^{\rho(t)}$ for $\rho(t) < r$ for some

$r \in \{1, \ldots, \eta\}$ and let $t = [t_1, \ldots, t_m]$ be any tree such that $\rho(t) = r$. Since $\rho(t_i) < r$ for $i = 1, \ldots, m$, $\phi(t_i) = c^{\rho(t_i)}$, whence

$$\phi(t) = \rho(t) A \prod_{i=1}^{m} \phi(t_i) = \rho(t) A \prod_{i=1}^{m} c^{\rho(t_i)} = \rho(t) A c^{\rho(t)-1} = c^{\rho(t)}$$

where we have used Theorem 2.4, $\sum_{i=1}^{m} \rho(t_i) = \rho(t) - 1$, $\rho(t) = r$ and the assumption that (2.4) holds for $r \in \{1, \ldots, \eta\}$. Therefore, (2.5) holds.

Conversely, suppose that (2.5) holds for some $\eta \geq 1$. Therefore, for any $r \in \{1, \ldots, \eta\}$, $\phi([\tau^{r-1}]) = c^r$, since $\rho([\tau^{r-1}]) = r$. On the other hand,

$$\phi([\tau^{r-1}]) = \rho([\tau^{r-1}]) A \left(\phi(\tau)\right)^{r-1} = rAc^{r-1}$$

where we have used Theorem 2.4, $\rho([\tau^{r-1}]) = r$ and $\phi(\tau) = c$. Therefore, $Ac^{r-1} = \frac{1}{r} c^r$, establishing that (2.4) holds. $\square$

We now give some new definitions and a few related results used in §3.

Definition 2.2

$$\begin{aligned}
\Phi_r &= \{\phi(t) : \forall\, t \in T \ s.t.\ \rho(t) = r\}. \\
\Omega_r &= \left\{ \prod_{j=1}^{\nu} \phi_j : \phi_j \in \Phi_{r_j},\ j = 1, \ldots, \nu,\ \sum_{j=1}^{\nu} r_j = r \right\}. \\
\Psi_r &= \left\{ \phi_1 + \sum_{j=1}^{\nu} \kappa_j (\phi_{j,1} - \phi_{j,2}) : \phi_1,\ \phi_{j,1},\ \phi_{j,2} \in \Phi_r,\ \kappa_j \in \mathbb{R},\ j = 1, \ldots, \nu \right\}. \\
\Upsilon_r &= \left\{ \phi_1 + \sum_{j=1}^{\nu} \kappa_j (\phi_{j,1} - \phi_{j,2}) : \phi_1,\ \phi_{j,1},\ \phi_{j,2} \in \Omega_r,\ \kappa_j \in \mathbb{R},\ j = 1, \ldots, \nu \right\}.
\end{aligned}$$

It follows immediately from these definitions that $\Phi_r \subset \Omega_r$, $\Phi_r \subset \Psi_r$, $\Phi_r \subset \Upsilon_r$, $\Omega_r \subset \Upsilon_r$ and $\Psi_r \subset \Upsilon_r$. Moreover, $\Phi_r = \Omega_r = \Psi_r = \Upsilon_r$ for $r \leq \eta$ if and only if $C(\eta)$ holds, since, by Theorem 2.5, $\phi(t) = c^{\rho(t)}$ for $\rho(t) \leq \eta$ if and only if $C(\eta)$ holds. Because $C(1)$ always holds, $\Phi_r = \Omega_r = \Psi_r = \Upsilon_r$ for $r = 0$ or 1. In addition, since $[\tau]$ is the only tree of order 2, $\Phi_2 = \Psi_2$.

The following result is also used in §3.

Lemma 2.6

(a) If $\phi_1 \in \Omega_{r_1}, \ldots, \phi_m \in \Omega_{r_m}$, *then*

$$rA \prod_{i=1}^{m} \phi_i \in \Phi_r \quad \textit{for} \quad r = 1 + \sum_{i=1}^{m} r_i.$$

(b) Similarly, if $\phi_1 \in \Upsilon_{r_1}, \ldots, \phi_m \in \Upsilon_{r_m}$, *then*

$$rA \prod_{i=1}^{m} \phi_i \in \Psi_r \quad \textit{for} \quad r = 1 + \sum_{i=1}^{m} r_i.$$

(c) *If the RK formula is of order p and $r \le p$, then*

$$\tilde\phi \in \tilde\Phi_r, \quad \tilde\phi \in \tilde\Omega_r, \quad \tilde\phi \in \tilde\Psi_r \quad or \quad \tilde\phi \in \tilde\Upsilon_r \quad \Rightarrow \quad \tilde\phi_{s+1} = 1.$$

Proof: To prove (a), note that, if $\phi_i \in \Omega_{r_i}$, then

$$\phi_i = \prod_{j=1}^{\nu_i} \phi(t_{i,j}) \quad \text{where each } \phi(t_{i,j}) \in \Phi_{\rho(t_{i,j})} \text{ and } r_i = \sum_{j=1}^{\nu_i} \rho(t_{i,j}).$$

Let $t = [t_{1,1}, \ldots, t_{1,\nu_1}, t_{2,1}, \ldots, t_{m,\nu_m}]$ and note $r = \rho(t)$. This and Theorem 2.4 imply that

$$rA \prod_{i=1}^{m} \phi_i = \rho(t) A \prod_{i=1}^{m} \left(\prod_{j=1}^{\nu_i} \phi(t_{i,j}) \right) = \phi(t) \in \Phi_r.$$

To prove (b), note

$$rA \prod_{i=1}^{m} \phi_i = rA \prod_{i=1}^{m} \left(\phi_{i,0} + \sum_{j=1}^{\nu_i} \kappa_{i,j} (\phi_{i,j,1} - \phi_{i,j,2}) \right) = rA \prod_{i=1}^{m} \phi_{i,0} + rA \sum_{l} \lambda_l$$

where $\phi_{i,0}$, $\phi_{i,j,1}$, $\phi_{i,j,2} \in \Upsilon_{r_i}$ and each $\lambda_l = \prod_{i=1}^{m} \beta_{l,i}$ with each $\beta_{l,i}$ being either a $\phi_{i,0}$ or a $\kappa_{i,j}(\phi_{i,j,1} - \phi_{i,j,2})$. At least one $\beta_{l,i}$, say $\beta_{l,k}$, must be of the latter form. Keeping this $\beta_{l,k}$ fixed and expanding the rest of the product, we see that λ_l becomes a sum of terms of the form

$$\bar\kappa_l \left(\prod_{\substack{i=1 \\ i \ne k}}^{m} \xi_{l,i} \right) (\xi_{l,k,1} - \xi_{l,k,2})$$

where $\xi_{l,i} \in \Omega_{r_i}$ and $\xi_{l,k,1}$, $\xi_{l,k,2} \in \Omega_{r_k}$. Hence,

$$\begin{aligned} rA \prod_{i=1}^{m} \phi_i &= rA \prod_{i=1}^{m} \phi_{i,0} + \sum_{l} \bar\kappa_l \, rA \left[\left(\prod_{\substack{i=1 \\ i \ne k}}^{m} \xi_{l,i} \right) (\xi_{l,k,1} - \xi_{l,k,2}) \right] \\ &= rA \prod_{i=1}^{m} \phi_{i,0} + \sum_{l} \bar\kappa_l \left\{ rA \left[\left(\prod_{\substack{i=1 \\ i \ne k}}^{m} \xi_{l,i} \right) \xi_{l,k,1} \right] - rA \left[\left(\prod_{\substack{i=1 \\ i \ne k}}^{m} \xi_{l,i} \right) \xi_{l,k,2} \right] \right\} \\ &= \hat\phi_1 + \sum_{l=1} \bar\kappa_l (\hat\phi_{l,1} - \hat\phi_{l,2}) \in \Psi_r \end{aligned}$$

since $\hat\phi_1$, $\hat\phi_{l,1}$, $\hat\phi_{l,2} \in \Phi_r$ by part (a).

To prove part (c), recall that, if $\tilde\phi \in \tilde\Phi_r$ and $r \le p$, then $\tilde\phi_{s+1} = 1$ by Theorem 2.4. It immediately follows that, if $\tilde\phi \in \tilde\Omega_r$ and $r \le p$, then $\tilde\phi_{s+1} = 1$ too. If $\tilde\phi \in \tilde\Psi_r$, then

$$\tilde\phi = \tilde\phi_1 + \sum_{j=1}^{\nu} \kappa_j (\tilde\phi_{j,1} - \tilde\phi_{j,2}) \tag{2.6}$$

for some $\tilde{\phi}_1$, $\tilde{\phi}_{j,1}$, $\tilde{\phi}_{j,2} \in \tilde{\Phi}_r$ and $\kappa_j \in \mathbb{R}$. If $r \leq p$, then $(\tilde{\phi}_1)_{s+1} = (\tilde{\phi}_{j,1})_{s+1} = (\tilde{\phi}_{j,2})_{s+1} = 1$, which together with (2.6) implies $\tilde{\phi}_{s+1} = 1$. A similar argument shows that, if $\phi \in \tilde{\Upsilon}_r$ and $r \leq p$, then $\tilde{\phi}_{s+1} = 1$. □

We end this section with the definition of three functions defined on trees, an example illustrating each, and three associated lemmas that are used in §3.

Definition 2.3 *For $q \geq 0$, the q*-lifted height *$H(t,q)$ is given by*

$$\begin{aligned}
H(t,q) &= 0 & &\forall\, t \in T \ s.t.\ \rho(t) \leq q,\\
H(t,q) &= 1 & &\forall\, t \in T \ s.t.\ \rho(t) = q+1,\\
H(t=[t_1,\ldots,t_m],q) &= 1 + \max_{i=1,\ldots,m} H(t_i,q) & &\forall\, t \in T \ s.t.\ \rho(t) \geq q+2.
\end{aligned}$$

$H(t,q)$ is a generalization of the *height* $H(t)$ of a tree t defined in [4, p. 88]: $H(t) = H(t,0)$.

Definition 2.4 *For $q \geq 0$, the q*-lifted number of ramifications *$R(t,q)$ is given by*

$$\begin{aligned}
R(t,q) &= 0, & &\forall\, t \in T \ s.t.\ \rho(t) \leq q,\\
R(t,q) &= 1, & &\forall\, t \in T \ s.t.\ \rho(t) = q+1,\\
R(t=[t_1,\ldots,t_m],q) &= \left\{ \begin{array}{ll} R(t_1,q) & \text{if } m=1 \\ 1 + \max_{i=1,\ldots,m} R(t_i,q) & \text{if } m \geq 2 \end{array} \right\} & &\forall\, t \in T \ s.t.\ \rho(t) \geq q+2.
\end{aligned}$$

Thus, for $\rho(t) \geq 1$, $R(t) = R(t,0)$ is 1 + the maximum number of ramifications along any branch of the tree.

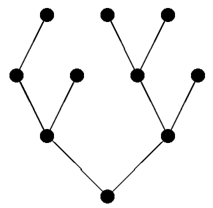
Definition 2.5 *For $q \geq r \geq 0$, the r,q*-lifted Newton doubling index *$D(t,r,q)$ is given by*

$D(t,r,q) = 0 \quad \forall\, t \in T$ *s.t.* $\rho(t) \leq q$,
$D(t,r,q) = 1 \quad \forall\, t \in T$ *s.t.* $\rho(t) = q+1$.

For $\rho(t) \geq q+2$, *let* $t = [t_1,\ldots,t_m]$.
If $m = 1$ *then*
  $D(t,r,q) = D(t_1,r,q)$
else
  let $\gamma_q = \max_{i=1,\ldots,m} D(t_i,r,q)$,
    choose t_R *such that* $D(t_R,r,q) = \gamma_q$ *and*
    let $\gamma_r = \max_{\substack{i=1,\ldots,m \\ i \neq R}} D(t_i,r,r)$
  if $\gamma_q \leq \gamma_r$ *then*
    $D(t,r,q) = \gamma_q + 1$
  else
    $D(t,r,q) = \gamma_q$
  end if
end if

We show below in Lemma 2.7 that $D(t, r, q) \leq D(t, r, r)$, whence, if there are two or more trees t_j and t_k, $j \neq k$, such that $\gamma_q = D(t_j, r, q) = D(t_k, r, q)$, then, no matter which of these trees one chooses for t_R, one always gets $\gamma_q \leq \gamma_r$ and so $D(t, r, q) = \gamma_q + 1$. Consequently, $D(t, r, q)$ is well-defined. Also note, for $\rho(t) \geq 1$, $D(t) = D(t, 0, 0)$ is 1 + the maximum number of times the number of branches in t is doubled.

Example 2.1 *For the tree*

$t = [[[\tau]\tau][[\tau\tau]\tau]] =$ 

$H(t) = 4$, $R(t) = 4$, $D(t) = 3$, $H(t, 2) = 3$, $R(t, 2) = 3$ and $D(t, 1, 2) = 2$.

Lemma 2.7 *If $q \geq r \geq 0$, then $D(t, q, q) \leq D(t, r, q) \leq D(t, r, r)$.*

Proof: The proof is by induction on $\rho(t)$. If $\rho(t) = 0$, then

$$D(t, q, q) = D(t, r, q) = D(t, r, r) = 0,$$

so the lemma holds for the base case.

For the induction step, let $\bar{\rho} \geq 1$ and assume that

$$D(t, q, q) \leq D(t, r, q) \leq D(t, r, r) \quad \forall\, t \in T \text{ s.t. } \rho(t) < \bar{\rho}. \tag{2.7}$$

Let $t = [t_1, \ldots, t_m]$ be any tree for which $\rho(t) = \bar{\rho}$. We show $D(t, q, q) \leq D(t, r, q)$ first. By the induction hypothesis (2.7), $D(t_i, q, q) \leq D(t_i, r, q) \leq D(t_i, r, r)$ for $i = 1, \ldots, m$ since $\rho(t_i) < \bar{\rho}$, whence

$$\max_{1 \leq i \leq m} D(t_i, q, q) \leq \max_{1 \leq i \leq m} D(t_i, r, q).$$

Furthermore, it follows immediately from Definition 2.5 that

$$D(t, q, q) = \max_{1 \leq i \leq m} D(t_i, q, q) \quad \text{or} \quad D(t, q, q) = 1 + \max_{1 \leq i \leq m} D(t_i, q, q)$$

and

$$D(t, r, q) = \max_{1 \leq i \leq m} D(t_i, r, q) \quad \text{or} \quad D(t, r, q) = 1 + \max_{1 \leq i \leq m} D(t_i, r, q).$$

To complete the proof that $D(t, q, q) \leq D(t, r, q)$, consider the following three cases.

1. If $D(t, q, q) = \max_i D(t_i, q, q)$, then

$$D(t, q, q) = \max_{1 \leq i \leq m} D(t_i, q, q) \leq \max_{1 \leq i \leq m} D(t_i, r, q) \leq D(t, r, q).$$

2. If $\max_i D(t_i, q, q) < \max_i D(t_i, r, q)$, then, since $D(t_i, q, q)$ and $D(t_i, r, q)$ are integers for $i = 1, \ldots, m$, $1 + \max_i D(t_i, q, q) \leq \max_i D(t_i, r, q)$, whence

$$D(t, q, q) \leq 1 + \max_{1 \leq i \leq m} D(t_i, q, q) \leq \max_{1 \leq i \leq m} D(t_i, r, q) \leq D(t, r, q).$$

3. If neither case 1 or 2 above is satisfied, then we must have that

$$D(t, q, q) = 1 + \max_{1 \leq i \leq m} D(t_i, q, q) \quad \text{and} \quad \max_i D(t_i, q, q) = \max_i D(t_i, r, q). \tag{2.8}$$

   The first of the relations (2.8) and Definition 2.5 imply that there must be at least two subtrees t_j and t_k of $t = [t_1, \ldots, t_m]$, $j \neq k$, such that

$$D(t_j, q, q) = D(t_k, q, q) = \max_{1 \leq i \leq m} D(t_i, q, q). \tag{2.9}$$

   By the second of the relations (2.8), (2.9) and the induction hypothesis (2.7)

$$\max_{1 \leq i \leq m} D(t_i, r, q) = \max_{1 \leq i \leq m} D(t_i, q, q) = D(t_j, q, q) \leq D(t_j, r, r)$$

   and

$$\max_{1 \leq i \leq m} D(t_i, r, q) = \max_{1 \leq i \leq m} D(t_i, q, q) = D(t_k, q, q) \leq D(t_k, r, r).$$

   Therefore, by Definition 2.5,

$$D(t, r, q) = 1 + \max_{1 \leq i \leq m} D(t_i, r, q). \tag{2.10}$$

   Combining (2.8) with (2.10), we get that

$$D(t, q, q) = 1 + \max_{1 \leq i \leq m} D(t_i, q, q) = 1 + \max_{1 \leq i \leq m} D(t_i, r, q) = D(t, r, q).$$

A similar argument shows that $D(t, r, q) \leq D(t, r, r)$. $\square$

Lemma 2.8 *If $q \geq r \geq 0$, $k \geq 1$, $t = [t_1, \ldots, t_m]$ and $D(t, r, q) \leq k$, then either*

1. *$D(t_i, r, q) \leq k - 1$ for $i = 1, \ldots, m$, or*
2. *$D(t_i, r, q) = k$ for exactly one $i \in \{1, \ldots, m\}$ and $D(t_j, r, q) \leq D(t_j, r, r) \leq k - 1$ for $j \neq i$, $j \in \{1, \ldots, m\}$.*

In case 2, we may renumber the subtrees $t_1, \ldots, t_m$ so that $D(t_i, r, q) \leq D(t_i, r, r) \leq k - 1$ for $i = 1, \ldots, m - 1$ and $D(t_m, r, q) = k$.

Proof: By Definition 2.5 and the assumptions of this lemma,

$$\max_{1 \le 1 \le m} D(t_i, r, q) \le D(t, r, q) \le k,$$

so $D(t_i, r, q) \le k$ for $i = 1, \ldots, m$. Therefore, either $D(t_i, r, q) \le k - 1$ for $i = 1, \ldots, m$ and case 1 holds or $D(t_i, r, q) = k$ for some $i \in \{1, \ldots, m\}$. In the latter case, $D(t_j, r, r) \le k-1$ for $i \ne j$, $j \in \{1, \ldots, m\}$, since otherwise we would have $D(t, r, q) = k + 1$ by Definition 2.5. To complete the proof of case 2, note $D(t_j, r, q) \le D(t_j, r, r)$ for $j \in \{1, \ldots, m\}$ by Lemma 2.7. To justify the final claim of the lemma, recall that the ordering of the subtrees $t_1, \ldots, t_m$ in $[t_1, \ldots, t_m]$ is immaterial. Therefore, we may renumber the subtrees $t_1, \ldots, t_m$ any way we choose. □

Lemma 2.9 *For $k \ge 1$,*

$$\begin{aligned} H(t, q) &= k, \quad \Rightarrow \quad \rho(t) \ge k + q, && (2.11) \\ R(t, q) &= k, \quad \Rightarrow \quad \rho(t) \ge 2k + q - 1, && (2.12) \\ D(t, r, q) &= k, \quad \Rightarrow \quad \rho(t) \ge 2^{k-1}(r + 2) + q - r - 1. && (2.13) \end{aligned}$$

Proof: Let $t_{H,k}$, $t_{R,k}$, $t_{D,q,k}$ and $t_{D,r,k}$ be (the generally nonunique) trees of minimal order satisfying $H(t_{H,k}, q) = k$, $R(t_{R,k}, q) = k$, $D(t_{D,q,k}, r, q) = k$ and $D(t_{D,r,k}, r, r) = k$, and let t_q and t_r be any trees of order $1 + q$ and $1 + r$, respectively. It follows immediately that $t_{H,1} = t_q$, $t_{R,1} = t_q$, $t_{D,q,1} = t_q$ and $t_{D,r,1} = t_r$ are possible choices for the trees $t_{H,k}$, $t_{R,k}$, $t_{D,q,k}$ and $t_{D,r,k}$ for $k = 1$. Moreover, $\rho(t_{H,1}) = \rho(t_{R,1}) = \rho(t_{D,q,1}) = 1 + q$ and $\rho(t_{D,r,1}) = 1 + r$, proving the result for $k = 1$. It is easy to show by induction on k that

$t_{H,k} = [t_{H,k-1}] = [_{k-1}t_q]_{k-1}$ and $\rho(t_{H,k}) = 1 + \rho(t_{H,k-1}) = k + q$;

$t_{R,k} = [\tau t_{R,k-1}] = ([\tau)_{k-1} t_q]_{k-1}$ and $\rho(t_{R,k}) = 2 + \rho(t_{R,k-1}) = 2k + q - 1$;

$t_{D,q,k} = [t_{D,q,k-1} t_{D,r,k-1}]$ and $\rho(t_{D,q,k}) = 1 + \rho(t_{D,q,k-1}) + \rho(t_{D,r,k-1}) = 2^{k-1}(r + 2) + q - r - 1$; and

$t_{D,r,k} = [t_{D,r,k-1} t_{D,r,k-1}]$ and $\rho(t_{D,r,k}) = 1 + 2\rho(t_{D,r,k-1}) = 2^{k-1}(r + 2) - 1$.

It also follows from this argument that $t_{D,q,k}$ is the tree formed from one copy of t_q and $2^{k-1} - 1$ copies of t_r by joining each pair of these trees to a new node, then joining each pair of these new nodes to yet another new node, etc. See Figure 2.1 for examples of the "minimal order trees" $t_{H,k}$, $t_{R,k}$ and $t_{D,q,k}$ for $k = 3$. □

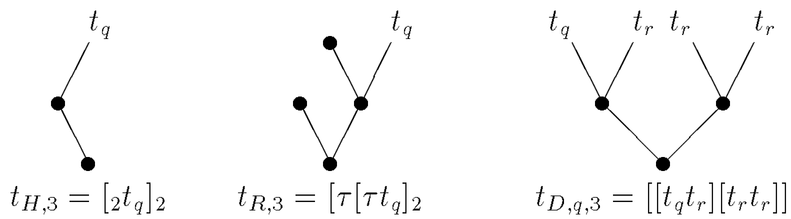


Figure 2.1: The "minimal order trees" $t_{H,k}$, $t_{R,k}$ and $t_{D,q,k}$ for $k = 3$.

3 The Proof of Theorem 1.1 and Related Results.

In this section, we prove Theorem 1.1 as well as several related results. These results can be used in a variety of other ways, one of which is to determine the error coefficients associated with $y_{n+1}^{(k)}$. We do not, though, pursue the latter analysis in detail in this paper.

To begin, rewrite the iteration (1.13)–(1.14) as

$$Y_n^{(k+1)} = y_n \varepsilon_s + hAf(Y_n^{(k)}) + hAJ_n^{(k)}(Y_n^{(k+1)} - Y_n^{(k)}) \tag{3.1}$$

$$y_{n+1}^{(k+1)} = y_n + hb^T f(Y_n^{(k)}) \tag{3.2}$$

where $k \geq 0$ is the iteration count. The three iteration schemes discussed in §1 differ from each other only in the choice of $J_n^{(k)}$.

To aid the discussion below, we introduce the *extended iterated internal stage vector* $\tilde{Y}_n^{(k)} = \left(Y_{n,1}^{(k)}, \ldots, Y_{n,s}^{(k)}, y_{n+1}^{(k)}\right)^T$, corresponding to $\tilde{Y}_n$ in (2.3), and the associated B-series

$$Y_n^{(k)} = B(\phi^{(k)}, y_n) \qquad \text{and} \qquad \tilde{Y}_n^{(k)} = B(\tilde{\phi}^{(k)}, y_n).$$

By Theorem 2.1, assumption (1.9) of Theorem 1.1 is equivalent to

$$\phi^{(0)}(t) = c^{\rho(t)} \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq q. \tag{3.3}$$

Similarly, conclusion (1.17) of Theorem 1.1 is equivalent to

$$\tilde{\phi}_{s+1}^{(k)}(t) = 1 \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq \min(p, q_k) \tag{3.4}$$

for $k \geq 1$. In the following three subsections, one for each case of Theorem 1.1, we investigate for which trees and under what conditions $\tilde{\phi}_{s+1}^{(k)}(t) = 1$.

3.1 Simple Iteration.

In this subsection, we consider the simple iteration (3.1)–(3.2) with $J_n^{(k)} = 0$, which we rewrite in the extended format as

$$\tilde{Y}_n^{(k+1)} = y_n \varepsilon_{s+1} + h\tilde{A}f(\tilde{Y}_n^{(k)}) \tag{3.1.1}$$

from which it follows that

Lemma 3.1.1 *If $k \geq 0$, then $\tilde{\phi}^{(k+1)}(\emptyset) = \varepsilon_{s+1}$. If in addition $\phi^{(k)}(\emptyset) = \varepsilon_s$, then, for $t = [t_1, \ldots, t_m]$ with $\rho(t) \geq 1$,*

$$\tilde{\phi}^{(k+1)}(t) = \rho(t)\tilde{A}\prod_{i=1}^{m}\tilde{\phi}^{(k)}(t_i) \tag{3.1.2}$$

and, in particular, $\tilde{\phi}^{(k+1)}(\tau) = \tilde{c}$.

Proof: Writing (3.1.1) as a B-series, we immediately see that $\tilde{\phi}^{(k+1)}(\emptyset) = \varepsilon_{s+1}$ for $k \geq 0$ independently of $\tilde{\phi}^{(k)}$, establishing the first statement of the lemma. To prove the second statement, assume $k \geq 0$ and $\tilde{\phi}^{(k)}(\emptyset) = \varepsilon_{s+1}$. By Theorem 2.2, (3.1.1) can be rewritten in B-series notation as

$$\sum_{t\in T} \tilde{\phi}^{(k+1)}(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} = y_n \varepsilon_{s+1} + \tilde{A} \sum_{t\in T} \tilde{\phi}'^{(k)}(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!}. \tag{3.1.3}$$

Note, though, that (3.1.3) holds under the weaker assumption that $\phi^{(k)}(\emptyset) = \varepsilon_s$, since the last column of $\tilde{A}$ is zero so the right side of (3.1.3) does not depend on $\tilde{\phi}^{(k)}_{s+1}(t)$ for any $t \in T$. Comparing corresponding trees in (3.1.3), we see

$$\tilde{\phi}^{(k+1)}(t) = \tilde{A}\tilde{\phi}'^{(k)}(t) = \rho(t)\tilde{A} \prod_{i=1}^{m} \tilde{\phi}^{(k)}(t_i) \quad \text{for } t = [t_1, \ldots, t_m] \text{ with } \rho(t) \geq 1$$

and, in particular, $\tilde{\phi}^{(k+1)}(\tau) = \tilde{A}\tilde{\phi}^{(k)}(\emptyset) = \tilde{A}\varepsilon_{s+1} = \tilde{c}$. □

Lemma 3.1.2 *For r, q, ν, k, $l \geq 0$,*

$$\begin{aligned} &\phi^{(k)}(t) = \phi(t)\ \forall\, t \in T \ s.t.\ H(t,r) \leq l \\ &\quad \Rightarrow \tilde{\phi}^{(k+1)}(t) = \tilde{\phi}(t)\ \forall\, t \in T \ s.t.\ H(t,r) \leq l+1, \end{aligned} \tag{3.1.4}$$

$$\begin{aligned} &\phi^{(k)}(t) \in \Omega_{\rho(t)}\ \forall\, t \in T \ s.t.\ H(t,q) \leq l \\ &\quad \Rightarrow \tilde{\phi}^{(k+1)}(t) \in \tilde{\Phi}_{\rho(t)}\ \forall\, t \in T \ s.t.\ H(t,q) \leq l+1, \end{aligned} \tag{3.1.5}$$

$$\begin{aligned} &\phi^{(k)}(t) \in \Upsilon_{\rho(t)}\ \forall\, t \in T \ s.t.\ H(t,\nu) \leq l \\ &\quad \Rightarrow \tilde{\phi}^{(k+1)}(t) \in \tilde{\Psi}_{\rho(t)}\ \forall\, t \in T \ s.t.\ H(t,\nu) \leq l+1. \end{aligned} \tag{3.1.6}$$

Moreover, if $\phi^{(k)}(\emptyset) = \varepsilon_s$ and A is nonsingular, then the $\Rightarrow$ in each of (3.1.4), (3.1.5) and (3.1.6) can be replaced by a $\Leftrightarrow$.

Proof: Consider the $\Rightarrow$ implications first. To begin assume $\rho(t) = 0$ and note $\tilde{\Phi}_0 = \tilde{\Psi}_0 = \{\tilde{\phi}(\emptyset)\}$, since $\emptyset$ is the only tree of order 0. By Lemma 3.1.1 and Theorem 2.4, $\tilde{\phi}^{(k+1)}(\emptyset) = \varepsilon_{s+1} = \tilde{\phi}(\emptyset)$. Hence the right side of each implication holds independently of $\phi^{(k)}$, proving all three implications in the $\Rightarrow$ direction for $\rho(t) = 0$. To complete the proof of (3.1.4) in the $\Rightarrow$ direction, note $\phi^{(k)}(\emptyset) = \varepsilon_s$ follows from the hypothesis (i.e., left side) of (3.1.4) since $l \geq 0$ and $H(\emptyset, r) = 0$. Therefore, we can apply Lemma 3.1.1. Let t be any tree satisfying $\rho(t) \geq 1$ and $H(t,r) \leq l+1$. From Definition 2.3, $t = [t_1, \ldots, t_m]$ with $H(t_i, r) \leq l$ for $i = 1, \ldots, m$, whence $\phi^{(k)}(t_i) = \phi(t_i)$ for $i = 1, \ldots, m$ by the hypothesis (i.e., left side) of (3.1.4). Therefore,

$$\tilde{\phi}^{(k+1)}(t) = \rho(t)\tilde{A} \prod_{i=1}^{m} \tilde{\phi}^{(k)}(t_i) = \rho(t)\tilde{A} \prod_{i=1}^{m} \tilde{\phi}(t_i) = \tilde{\phi}(t),$$

where the first and third equalities follow from Lemma 3.1.1 and Theorem 2.4, respectively, and the middle one follows from $\prod \phi^{(k)}(t_i) = \prod \phi(t_i)$ and the observation that the last column of $\tilde{A}$ is zero. Implications (3.1.5) and (3.1.6) in the $\Rightarrow$ direction follow from a similar argument, but using Lemma 2.6 instead of Theorem 2.4.

Assume $\phi^{(k)}(\emptyset) = \varepsilon_s$ and A is nonsingular. To prove (3.1.4) in the $\Leftarrow$ direction, let t be any tree satisfying $H(t, r) \leq l$, whence $H([t], r) \leq l + 1$ and $\phi^{(k+1)}([t]) = \phi([t])$ by the hypothesis of (3.1.4) in the $\Leftarrow$ direction (i.e., right side of (3.1.4)). Consequently, $\rho([t])A\phi^{(k)}(t) = \rho([t])A\phi(t)$ by Lemma 3.1.1 and Theorem 2.4, from which it follows that $\phi^{(k)}(t) = \phi(t)$, since $\rho([t]) \neq 0$ and A is nonsingular, thus proving (3.1.4) in the $\Leftarrow$ direction. Similar arguments show that (3.1.5) and (3.1.6) hold in the $\Leftarrow$ direction also. $\square$

Theorem 3.1.3 *If r, q, $\nu \geq 0$ and $k \geq 1$, then*

$$\begin{aligned} &\phi^{(0)}(t) = \phi(t)\ \forall\, t \in T \ s.t.\ \rho(t) \leq r \\ &\qquad \Rightarrow \tilde{\phi}^{(k)}(t) = \tilde{\phi}(t)\ \forall\, t \in T\ s.t.\ H(t, r) \leq k, \end{aligned} \tag{3.1.7}$$

$$\begin{aligned} &\phi^{(0)}(t) \in \Omega_{\rho(t)}\ \forall\, t \in T\ s.t.\ \rho(t) \leq q \\ &\qquad \Rightarrow \tilde{\phi}^{(k)}(t) \in \tilde{\Phi}_{\rho(t)}\ \forall\, t \in T\ s.t.\ H(t, q) \leq k, \end{aligned} \tag{3.1.8}$$

$$\begin{aligned} &\phi^{(0)}(t) \in \Upsilon_{\rho(t)}\ \forall\, t \in T\ s.t.\ \rho(t) \leq \nu \\ &\qquad \Rightarrow \tilde{\phi}^{(k)}(t) \in \tilde{\Psi}_{\rho(t)}\ \forall\, t \in T\ s.t.\ H(t, \nu) \leq k. \end{aligned} \tag{3.1.9}$$

For the special case $r\,(= q = \nu) = -1$ and $k \geq 1$,

$$\tilde{\phi}^{(k)}(t) = \tilde{\phi}(t)\ \forall\, t \in T\ s.t.\ H(t) \leq k - 1. \tag{3.1.10}$$

Moreover, if $\phi^{(0)}(\emptyset) = \varepsilon_s$ and A is nonsingular, then the $\Rightarrow$ in each of (3.1.7), (3.1.8) and (3.1.9) can be replaced by a $\Leftrightarrow$. In addition, if $\phi^{(1)}(\tau) \neq \phi(\tau)$ and A is nonsingular, (3.1.10) is tight in the sense that $\phi^{(k)}([_k\emptyset]_k) \neq \phi([_k\emptyset]_k)$ for $k \geq 1$, where $H([_k\emptyset]_k) = k$.

Proof: Since $H(t, \mu) = 0$ if and only if $\rho(t) \leq \mu$, the $\Rightarrow$ direction for (3.1.7), (3.1.8) or (3.1.9) follows immediately from Lemma 3.1.2 with $l = k$ by induction on k. The $\Leftarrow$ direction follows by induction on k from a slight modification of Lemma 3.1.2 that removes the ˜s from the right side of (3.1.4), (3.1.5) and (3.1.6). For $r = -1$, $\tilde{\phi}^{(1)}(\emptyset) = \varepsilon_{s+1} = \tilde{\phi}(\emptyset)$ by Lemma 3.1.1 and Theorem 2.4. Therefore, (3.1.10) follows from Lemma 3.1.2 with $l = k - 1$ by induction on k. Similarly, recalling that $\tau = [\emptyset]$, one can easily show by induction on k that $\phi^{(k)}([_k\emptyset]_k) \neq \phi([_k\emptyset]_k)$ for $k \geq 1$ if A is nonsingular and $\phi^{(1)}(\tau) \neq \phi(\tau)$. $\square$

We need the assumption $\phi^{(1)}(\tau) \neq \phi(\tau)$ in the last part of Theorem 3.1.3 since we cannot apply any of the B-series theory in this subsection to draw this conclusion if $\phi^{(0)}(\emptyset) \neq \phi(\emptyset)$.

Proof of Case 1 of Theorem 1.1. As noted at the start of §3, the assumption of Theorem 1.1 that the initial guess $Y_n^{(0)}$ satisfies (1.9) is equivalent to (3.3), which in turn implies that the left side of (3.1.8) holds, since $c^{\rho(t)} \in \Omega_{\rho(t)}$ for all $t \in T$. Hence, if $q \geq 0$, then the right side of (3.1.8) holds. Since any $\tilde{\phi} \in \tilde{\Phi}_\mu$ satisfies $\tilde{\phi}_{s+1} = 1$ for $\mu \leq p$ by Theorem 2.4 and since $H(t, q) \geq k + 1$ implies $\rho(t) \geq q + k + 1 = q_k + 1$ by Lemma 2.9, it follows from the right side of (3.1.8) that $\tilde{\phi}_{s+1}^{(k)}(t) = 1$ for all t satisfying $\rho(t) \leq \min(p, q_k)$, which together with Theorem 2.1 implies Case 1 of Theorem 1.1 for $q \geq 0$. For $q = -1$, a similar argument using (3.1.10) instead of (3.1.8) establishes the result. $\square$

Case 1 of Theorem 1.1 can be sharpened a little by introducing the H *restricted order values*

$$p_k^H = \max\{\mu : \tilde{\phi}_{s+1}(t) = 1 \;\forall\, t \in T \text{ s.t. } \rho(t) \leq \mu \text{ and } H(t,1) \leq k\}. \tag{3.1.11}$$

For $q \geq 0$, if we know $\phi^{(0)}(t) = c^{\rho(t)}$ for $\rho(t) \leq q$ (and not just $\phi^{(0)}(t) \in \Omega_{\rho(t)}$ for $\rho(t) \leq q$), then it follows easily by induction on $k \geq 1$ that, for any $t \in T$ such that $H(t,q) \leq k$, there is an associated $\hat{t} \in T$ such that $\rho(t) = \rho(\hat{t})$, $H(\hat{t},1) \leq k$ and $\tilde{\phi}^{(k)}(t) = \tilde{\phi}(\hat{t})$, which of course implies the right side of (3.1.8). We therefore have that $\tilde{\phi}_{s+1}^{(k)}(t) = \tilde{\phi}_{s+1}(\hat{t})$ for $\rho(t) \leq q + k = q_k$ and $\tilde{\phi}_{s+1}(\hat{t}) = 1$ for $\rho(t) = \rho(\hat{t}) \leq p_k^H$, since $H(\hat{t},1) \leq k$. Consequently,

$$y_{n+1}^{(k)} = y(x_{n+1}; x_n, y_n) + O(h^{\min(p_k^H, q_k)+1}), \tag{3.1.12}$$

which is a slight improvement over (1.17), since $p_1^H \geq p_2^H \geq \cdots \geq p_p^H = p_{p+1}^H = \cdots = p$ and, for some RK methods, $p_k^H > p_{k+1}^H$ for some $k < p$. This result indicates that the order of $y_{n+1}^{(k)}$ may decrease as k increases, as illustrated in Example 4.5 of §4. For $q = -1$, (3.1.12) gives no improvement over (1.17), since the restriction inherent in (3.1.11) does not exclude any trees that are not already excluded in the basic argument used to prove (1.17).

It is natural to ask whether the results established in this subsection are tight. It is for this reason that we proved the $\Leftrightarrow$ implications in Lemma 3.1.2 and Theorem 3.1.3: they show that the $\Rightarrow$ implications are tight if A is nonsingular and either $\phi^{(k)}(\emptyset) = \varepsilon_s$ or $\phi^{(0)}(\emptyset) = \varepsilon_s$, respectively. These assumptions on $\phi^{(k)}(\emptyset)$ and $\phi^{(0)}(\emptyset)$ are nearly always satisfied in practice and are needed in our analysis only to ensure that the B-series results of Lemma 3.1.1 are applicable.

The $\Rightarrow$ implications are often tight even if A is singular, although examples can be constructed for which they are not. In particular, if the underlying RK formula is explicit, whence A is strictly lower triangular, then it is easy to show that, if $q \geq 0$, then $\tilde{\phi}^{(k)}(t) = 0 = \tilde{\phi}(t)$ for $H(t) \geq s+1$ and $k \geq s+1$. In addition, from (3.1.7), $\tilde{\phi}^{(k)}(t) = \tilde{\phi}(t)$ for $H(t) \leq k$. Therefore, if $k \geq s+1$, $\tilde{\phi}^{(k)}(t) = \tilde{\phi}(t)$ for all $t \in T$. If $q = -1$, then a similar argument shows that $\tilde{\phi}^{(k)}(t) = \tilde{\phi}(t)$ for all $t \in T$ if $k \geq s+2$. Thus, (3.1.4)–(3.1.6) and (3.1.7)–(3.1.10) are certainly not tight in either case. We also see in this example that, if either $q \geq 0$ and $k \geq s+1$ or $q = -1$ and $k \geq s+2$, then $y_{n+1}^{(k)} = y_{n+1}$, an unusual occurrence for a nonlinear fixed-point iteration.

Since the proof of Case 1 of Theorem 1.1 is based largely on (3.1.8), one might initially guess that the former result (or the corresponding result with (1.17) replaced by (3.1.12)) would be tight too for A nonsingular. But a little reflection shows this is not the case, since it is easy to construct examples for which $\phi^{(0)}(t) \notin \Omega_{\rho(t)}$ but $\tilde{\phi}_{s+1}^{(k)}([_k t]_k) = \rho([_k t]_k) \cdots \rho([t]) b^T A^{k-1} \phi^{(0)}(t) = 1$ for some $k \geq 1$. Or to put it more concretely, examples exist for which q is the largest value such that the initial guess $Y_n^{(0)}$ satisfies (1.9), but

$$y_{n+1}^{(k)} = y(x_{n+1}; x_n, y_n) + O(h^{\mu+1}) \tag{3.1.13}$$

for some $k \geq 1$ and some $\mu > \min(p_k^H, q_k) \geq \min(p, q_k)$. See for instance Example 2 of [11]. However, the B-series for the initial guess $Y_n^{(0)}$ often satisfies the stronger assumption

$$\phi^{(0)}(t) = \zeta(t)\, c^{\rho(t)} \quad \forall t \in T, \quad \text{where } \zeta : T \to \mathbb{R} \text{ and } \zeta(t) = 1 \text{ for } \rho(t) \leq q. \tag{3.1.14}$$

This stronger assumption holds, for example, if the same RK starter formula is used to calculate each component of $Y_n^{(0)}$, where, for each $i \in \{1, \ldots, s\}$, the scaled stepsize $h_i = c_i h$ is used in the computation for $Y_{n,i}^{(0)}$. In this case, we get the tight result

Theorem 3.1.4 *Let the B-series for the initial guess $Y_n^{(0)}$ satisfy (3.1.14) and assume $q \geq 0$ is chosen as large as possible. If either $q \geq 1$ or $q_k < p_k^H$, then the order of $y_{n+1}^{(k)}$ is exactly $\min(p_k^H, q_k)$, where p_k^H is defined by (3.1.11) and $q_k = q + k$.*

Proof: From the discussion following (3.1.11), we know that the order of $y_{n+1}^{(k)}$ is at least $\min(p_k^H, q_k)$. Therefore, we need only show that (3.1.13) does not hold for any $\mu > \min(p_k^H, q_k)$ in general.

First assume $q_k < p_k^H$. From the hypothesis of the theorem, there is a tree $\hat{t}$ such that $\rho(\hat{t}) = q + 1$ and $\zeta(\hat{t}) \neq 1$. Therefore, by Lemma 3.1.1 and Theorem 2.4,

$$\begin{aligned}
\tilde{\phi}_{s+1}^{(k)}([_k\hat{t}]_k) &= (q+k+1) b^T \phi^{(k-1)}([_{k-1}\hat{t}]_{k-1}) \\
&= (q+k+1)\cdots(q+2) b^T A^{k-1} \phi^{(0)}(\hat{t}) \\
&= \zeta(\hat{t})(q+k+1)\cdots(q+2) b^T A^{k-1} c^{q+1} \\
&= \zeta(\hat{t})\tilde{\phi}_{s+1}([_k\tau^{q+1}]_k) \\
&= \zeta(\hat{t}) \neq 1,
\end{aligned}$$

where we have used $\tilde{\phi}_{s+1}([_k\tau^{q+1}]_k) = 1$, which follows from $\rho([_k\tau^{q+1}]_k) = q+k+1 = q_k+1 \leq p_k^H$, $H([_k\tau^{q+1}]_k, 1) = k$, and the definition of p_k^H. Since $\tilde{\phi}_{s+1}^{(k)}([_k\hat{t}]_k) \neq 1$ for $\rho([_k\hat{t}]_k) = q+k+1 = q_k+1$, the order of $y_{n+1}^{(k)}$ is at most $q_k = \min(p_k^H, q_k)$ in general, completing the proof for this case.

Next assume $q \geq 1$. Since we have already proven the result for $q_k < p_k^H$, we need only consider $q \geq 1$ and $p_k^H \leq q_k$. In this case, there exists a tree $\hat{t}$ such that $\rho(\hat{t}) = p_k^H + 1$, $H(\hat{t}, 1) \leq k$ and $\tilde{\phi}_{s+1}(\hat{t}) \neq 1$. Moreover, since $q \geq 1$, $\phi^{(0)}(t) = c^{\rho(t)} = \phi(t)$ for $\rho(t) \leq 1$, whence we may take $r = 1$ in (3.1.7) to conclude that $\tilde{\phi}^{(k)}(t) = \tilde{\phi}(t)$ for all $t \in T$ such that $H(t, 1) \leq k$. Therefore, $\tilde{\phi}_{s+1}^{(k)}(\hat{t}) = \tilde{\phi}_{s+1}(\hat{t}) \neq 1$, whence the order of $y_{n+1}^{(k)}$ is at most $p_k^H = \min(p_k^H, q_k)$ in general. $\square$

The rather odd restriction that either $q \geq 1$ or $q_k < p_k^H$ in Theorem 3.1.4 can not be relaxed, since it is easy to find examples for which $q = 0$, $p_1^H = q_1 = 1$ and the order of $y_{n+1}^{(1)}$ is greater than 1.

3.2 The Modified Newton Iteration.

In this subsection, we consider the modified Newton iteration, (3.1)–(3.2) with

$$J_n^{(k)} = \mathrm{diag}(f'(y_n), \ldots, f'(y_n)).$$

It follows that the B-series for $\tilde{Y}_n^{(k+1)}$ satisfies

Lemma 3.2.1 *If $k \geq 0$, then $\tilde{\phi}^{(k+1)}(\emptyset) = \varepsilon_{s+1}$. If in addition $\phi^{(k)}(\emptyset) = \varepsilon_s$, then, for $t = [t_1, \ldots, t_m]$ with $\rho(t) \geq 1$,*

$$\phi^{(k+1)}(t) = \begin{cases} \rho(t) A \phi^{(k+1)}(t_1) & \text{if } m = 1 \text{ and } \rho(t_1) \geq 0, \\ \rho(t) A \prod_{i=1}^{m} \phi^{(k)}(t_i) & \text{if } m \geq 2 \text{ and } \rho(t_i) \geq 1 \text{ for } i = 1, \ldots, m, \end{cases} \tag{3.2.1}$$

$$\tilde{\phi}_{s+1}^{(k+1)}(t) = \rho(t) b^T \prod_{i=1}^{m} \phi^{(k)}(t_i) \tag{3.2.2}$$

and, in particular, $\tilde{\phi}^{(k+1)}(\tau) = \tilde{c}$.

Proof: Writing (3.1)–(3.2) as a B-series, we immediately see that $\tilde{\phi}^{(k+1)}(\emptyset) = \varepsilon_{s+1}$ for $k \geq 0$, independently of $\phi^{(k)}$, proving the first statement of the lemma. To prove the second statement, assume $k \geq 0$ and $\phi^{(k)}(\emptyset) = \varepsilon_s$. By Theorem 2.2, (3.1) can be rewritten in B-series notation as

$$\begin{aligned} \sum_{t \in T} \phi^{(k+1)}(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} &= y_n \varepsilon_s + A \sum_{t \in T} \phi'^{(k)}(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} \\ &+ hAf'(y_n) \sum_{t \in T} \left(\phi^{(k+1)}(t) - \phi^{(k)}(t) \right) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} \end{aligned} \tag{3.2.3}$$

where $\phi'^{(k)}(t) = \rho(t) \prod_{i=1}^{m} \phi^{(k)}(t_i)$ for $t = [t_1, \ldots, t_m]$. By (2.1), $f'(y_n) F(t)(y_n) = F([t])(y_n)$, whence the last sum in (3.2.3) can be rewritten as

$$\begin{aligned} & hAf'(y_n) \sum_{t \in T} \left(\phi^{(k+1)}(t) - \phi^{(k)}(t) \right) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} \\ &= \sum_{t \in T} (\rho(t) + 1) A \left(\phi^{(k+1)}(t) - \phi^{(k)}(t) \right) F([t])(y_n) \frac{h^{\rho(t)+1}}{(\rho(t)+1)!} \\ &= \sum_{t \in T} \rho([t]) A \left(\phi^{(k+1)}(t) - \phi^{(k)}(t) \right) F([t])(y_n) \frac{h^{\rho([t])}}{\rho([t])!}. \end{aligned} \tag{3.2.4}$$

Therefore, if $t = [t_1]$, (3.2.3) and (3.2.4) imply

$$\phi^{(k+1)}(t) = \rho(t) A \phi^{(k)}(t_1) + \rho(t) A \left(\phi^{(k+1)}(t_1) - \phi^{(k)}(t_1) \right) = \rho(t) A \phi^{(k+1)}(t_1)$$

which proves the first part of (3.2.1). To prove the second part, let $t = [t_1, \ldots, t_m]$ with $m \geq 2$ and $\rho(t_i) \geq 1$ for $i = 1, \ldots, m$. In this case, (3.2.4) contributes nothing to (3.2.3), whence

$$\phi^{(k+1)}(t) = A \phi'^{(k)}(t) = \rho(t) A \prod_{i=1}^{m} \phi^{(k)}(t_i)$$

proving the second part of (3.2.1). Since (3.2) is the same for simple iteration as for the modified Newton iteration, it follows from an argument similar to the one used in the proof of Lemma 3.1.1 that (3.2.2) holds. To verify the last equation of the lemma, recall that $\phi^{(k)}(\emptyset) = \phi^{(k+1)}(\emptyset) = \varepsilon_s$. Therefore, since $\tau = [\emptyset]$, it follows from the first part of (3.2.1) that

$$\phi^{(k+1)}(\tau) = A\phi^{(k+1)}(\emptyset) = A\varepsilon_s = c$$

and from (3.2.2) that

$$\phi_{s+1}^{(k+1)}(\tau) = b^T\phi^{(k)}(\emptyset) = b^T\varepsilon_s = c_{s+1}.$$

Combining these equations, we get $\tilde{\phi}^{(k+1)}(\tau) = \tilde{c}$, as required. $\square$

Lemma 3.2.2 *For r, q, ν, k, $l \geq 0$,*

$$\begin{aligned} &\phi^{(k)}(t) = \phi(t)\ \forall\, t \in T\ s.t.\ R(t,r) \leq l \\ &\qquad \Rightarrow \phi^{(k+1)}(t) = \phi(t)\ \forall\, t \in T\ s.t.\ R(t,r) \leq l+1, \end{aligned} \tag{3.2.5}$$

$$\begin{aligned} &\phi^{(k)}(t) \in \Omega_{\rho(t)}\ \forall\, t \in T\ s.t.\ R(t,q) \leq l \\ &\qquad \Rightarrow \phi^{(k+1)}(t) \in \Phi_{\rho(t)}\ \forall\, t \in T\ s.t.\ R(t,q) \leq l+1, \end{aligned} \tag{3.2.6}$$

$$\begin{aligned} &\phi^{(k)}(t) \in \Upsilon_{\rho(t)}\ \forall\, t \in T\ s.t.\ R(t,\nu) \leq l \\ &\qquad \Rightarrow \phi^{(k+1)}(t) \in \Psi_{\rho(t)}\ \forall\, t \in T\ s.t.\ R(t,\nu) \leq l+1. \end{aligned} \tag{3.2.7}$$

Moreover, if $\phi^{(k)}(\emptyset) = \varepsilon_s$, $\phi^{(k)}(\tau) = c$, $c_i \neq 0$ for $i = 1, \ldots, s$, and A is nonsingular, then the $\Rightarrow$ in (3.2.5) can be replaced by a $\Leftrightarrow$.

Proof: Consider the $\Rightarrow$ implications first, and to begin assume $\rho(t) = 0$. Note $\Phi_0 = \Psi_0 = \{\phi(\emptyset)\}$, since $\emptyset$ is the only tree of order 0. By Lemma 3.2.1 and Theorem 2.4, $\phi^{(k+1)}(\emptyset) = \varepsilon_s = \phi(\emptyset)$. Hence the right side of each implication holds independently of $\phi^{(k)}$, proving all three implications in the $\Rightarrow$ direction for $\rho(t) = 0$. To complete the proof of (3.2.5) in the $\Rightarrow$ direction, note $\phi^{(k)}(\emptyset) = \varepsilon_s$ follows from the hypothesis (i.e., left side) of (3.2.5), since $l \geq 0$ and $r \geq 0$ implies $R(\emptyset, r) = 0 \leq l$. Therefore, we can apply Lemma 3.2.1. Let t be any tree satisfying $\rho(t) \geq 1$ and $R(t,r) \leq l+1$. From Definition 2.4, either $t = [t_1]$ with $R(t_1, r) \leq l+1$ or $t = [t_1, \ldots, t_m]$ with $m \geq 2$, $\rho(t_i) \geq 1$ and $R(t_i, r) \leq l$ for $i = 1, \ldots, m$. If $t = [t_1]$, it follows easily by induction on $\rho(t)$ that $\phi^{(k+1)}(t) = \phi(t)$, since $\phi^{(k+1)}(\emptyset) = \varepsilon_s = \phi(\emptyset)$ and $\phi^{(k+1)}(t) = \rho(t)A\phi^{(k+1)}(t_1)$ by Lemma 3.2.1. Therefore, assume $t = [t_1, \ldots, t_m]$ with $m \geq 2$, $\rho(t_i) \geq 1$ and $R(t_i, r) \leq l$ for $i = 1, \ldots, m$. Hence, $\phi^{(k)}(t_i) = \phi(t_i)$ for $i = 1, \ldots, m$ by the hypothesis (i.e., left side) of (3.2.5). Consequently,

$$\phi^{(k+1)}(t) = \rho(t)A\prod_{i=1}^{m}\phi^{(k)}(t_i) = \rho(t)A\prod_{i=1}^{m}\phi(t_i) = \phi(t),$$

where the first and third equalities follow from Lemma 3.2.1 and Theorem 2.4, respectively, and the middle one follows from $\phi^{(k)}(t_i) = \phi(t_i)$ for $i = 1, \ldots, m$. Implications (3.2.6) and (3.2.7) in the $\Rightarrow$ direction follow from a similar argument, but using Lemma 2.6 instead of Theorem 2.4.

Assume $\phi^{(k)}(\emptyset) = \varepsilon_s$, $\phi^{(k)}(\tau) = c$, $c_i \neq 0$ for $i = 1, \ldots, s$, and A is nonsingular. It follows immediately from the first two of these assumptions and Theorem 2.4 that $\phi^{(k)}(\emptyset) = \phi(\emptyset)$ and $\phi^{(k)}(\tau) = \phi(\tau)$. To complete the proof of (3.2.5) in the $\Leftarrow$ direction, let t be any tree satisfying $R(t, r) \leq l$ and $\rho(t) \geq 2$. Since $R([\tau, t], r) \leq l + 1$, it follows from the right side of (3.2.5) that $\phi^{(k+1)}([\tau, t]) = \phi([\tau, t])$. Consequently, $\rho([\tau, t])A\phi^{(k)}(\tau)\phi^{(k)}(t) = \rho([\tau, t])A\phi(\tau)\phi(t)$ by Lemma 3.2.1 and Theorem 2.4, whence $\phi^{(k)}(\tau)\phi^{(k)}(t) = \phi(\tau)\phi(t)$, since $\rho([\tau, t]) \neq 0$ and A is nonsingular. Therefore, $\phi^{(k)}(t) = \phi(t)$ since $\phi^{(k)}(\tau) = c = \phi(\tau)$ and $c_i \neq 0$ for $i = 1, \ldots, s$. $\square$

Theorem 3.2.3 *If r, q, $\nu \geq 0$ and $k \geq 1$, then*

$$\begin{aligned} &\phi^{(0)}(t) = \phi(t)\ \forall\, t \in T \ s.t.\ \rho(t) \leq r \\ &\qquad \Rightarrow \phi^{(k)}(t) = \phi(t)\ \forall\, t \in T \ s.t.\ R(t, r) \leq k, \end{aligned} \tag{3.2.8}$$

$$\begin{aligned} &\phi^{(0)}(t) \in \Omega_{\rho(t)}\ \forall\, t \in T \ s.t.\ \rho(t) \leq q \\ &\qquad \Rightarrow \phi^{(k)}(t) \in \Phi_{\rho(t)}\ \forall\, t \in T \ s.t.\ R(t, q) \leq k, \end{aligned} \tag{3.2.9}$$

$$\begin{aligned} &\phi^{(0)}(t) \in \Upsilon_{\rho(t)}\ \forall\, t \in T \ s.t.\ \rho(t) \leq \nu \\ &\qquad \Rightarrow \phi^{(k)}(t) \in \Psi_{\rho(t)}\ \forall\, t \in T \ s.t.\ R(t, \nu) \leq k. \end{aligned} \tag{3.2.10}$$

For the special case $r\,(= q = \nu) = -1$ and $k \geq 1$,

$$\phi^{(k)}(t) = \phi(t)\ \forall\, t \in T \ s.t.\ R(t) \leq k - 1. \tag{3.2.11}$$

Moreover, if $\phi^{(0)}(\emptyset) = \varepsilon_s$, $\phi^{(0)}(\tau) = c$, $c_i \neq 0$ for $i = 1, \ldots, s$, and A is nonsingular, then the $\Rightarrow$ in (3.2.8) can be replaced by a $\Leftrightarrow$. In addition, if $(\phi^{(1)}(\tau))^2 \neq c^2$, $c_i \neq 0$ for $i = 1, \ldots, s$, and A is nonsingular, then (3.2.11) is tight in the sense that $\phi^{(k)}(t_{R,k}) \neq \phi(t_{R,k})$ for $k \geq 1$, where $t_{R,k}$ is defined recursively by $t_{R,1} = \tau$ and $t_{R,k+1} = [\tau, t_{R,k}]$ for $k \geq 1$, whence $R(t_{R,k}) = k$.

Proof: Since $\rho(t) \leq \mu$ if and only if $R(t, \mu) = 0$, the $\Rightarrow$ direction for (3.2.8), (3.2.9) or (3.2.10) follows immediately by induction on k from Lemma 3.2.2 with $l = k$. The $\Leftarrow$ direction for (3.2.8) also follows by induction on k from Lemma 3.2.2 with $l = k$. For $r = -1$, $\tilde{\phi}^{(1)}(\emptyset) = \varepsilon_{s+1} = \tilde{\phi}(\emptyset)$ by Lemma 3.2.1 and Theorem 2.4. Therefore, (3.2.11) follows by induction on k from Lemma 3.2.2 with $l = k - 1$.

To see that (3.2.11) is tight if $(\phi^{(1)}(\tau))^2 \neq c^2$, $c_i \neq 0$ for $i = 1, \ldots, s$, and A is nonsingular, first note that $\phi^{(1)}(\tau) \neq c = \phi(\tau)$, whence $\phi^{(k)}(t_{R,k}) \neq \phi(t_{R,k})$ holds for $k = 1$ since $t_{R,1} = \tau$ by definition. We repeatedly use Lemma 3.2.1 in the remainder of this proof. To begin, note that this lemma ensures that $\phi^{(k)}(\emptyset) = \varepsilon_s$ for $k \geq 1$, from which it follows that

$$\phi^{(2)}(t_{R,2}) = \rho(t_{R,2})A\left(\phi^{(1)}(\tau)\right)^2 \neq \rho(t_{R,2})Ac^2 = \phi(t_{R,2}),$$

since $(\phi^{(1)}(\tau))^2 \neq c^2$, $\rho(t_{R,2}) > 0$ and A is nonsingular. Moreover, since $\phi^{(k)}(\tau) = c$ for $k \geq 2$, it follows by induction on $k \geq 3$ that

$$\begin{aligned} \phi^{(k)}(t_{R,k}) &= \rho(t_{R,k})A\phi^{(k-1)}(\tau)\phi^{(k-1)}(t_{R,k-1}) \\ &= \rho(t_{R,k})Ac\phi^{(k-1)}(t_{R,k-1}) \\ &\neq \rho(t_{R,k})Ac\phi(t_{R,k-1}) = \phi(t_{R,k}), \end{aligned}$$

where we have used the induction hypothesis $\phi^{(k-1)}(t_{R,k-1}) \neq \phi(t_{R,k-1})$, the assumptions that $c_i \neq 0$ for $i = 1, \ldots, s$ and A is nonsingular as well as the observation that $\rho(t_{R,k}) > 0$. $\square$

Proof of Case 2 of Theorem 1.1. Assume first that $q \geq 0$. As noted at the start of §3, the assumption of Theorem 1.1 that the initial guess $Y_n^{(0)}$ satisfies (1.9) is equivalent to (3.3). That is, we may assume

$$\phi^{(0)}(t) = c^{\rho(t)} \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq q. \tag{3.2.12}$$

In particular, $\phi^{(0)}(\emptyset) = \varepsilon_s$. In addition, Lemma 3.2.1 ensures that $\tilde{\phi}^{(k)}(\emptyset) = \varepsilon_{s+1}$ for $k \geq 1$, whence $\tilde{\phi}^{(k)}_{s+1}(\emptyset) = 1$ for $k \geq 1$. Moreover, it follows from Lemma 3.2.1 that

$$\tilde{\phi}^{(k+1)}_{s+1}(t) = \rho(t) b^T \prod_{i=1}^{m} \phi^{(k)}(t_i) \tag{3.2.13}$$

for $t = [t_1, \ldots, t_m]$ and $k \geq 0$. We will now show that, for $k \geq 0$,

$$\tilde{\phi}^{(k+1)}_{s+1}(t) \in (\tilde{\Phi}_{\rho(t)})_{s+1} \quad \text{for } t = [t_1, \ldots, t_m] \text{ with } R(t_i, q) \leq k. \tag{3.2.14}$$

To this end, first note that (3.2.12) and Definition 2.4 ensure that

$$\phi^{(0)}(t_i) = c^{\rho(t_i)} \quad \forall\, t \in T \text{ s.t. } R(t_i, q) = 0. \tag{3.2.15}$$

Using (3.2.15), (3.2.13) for $k = 0$ and an argument similar to that used to prove Lemma 3.1.2, we can show (3.2.14) holds for $k = 0$. To prove (3.2.14) for $k \geq 1$, first note that (3.2.12) ensures that the left side of (3.2.9) holds, since $c^{\rho(t)} \in \Omega_{\rho(t)}$ for all $t \in T$. Therefore, by (3.2.9),

$$\phi^{(k)}(t_i) \in \Phi_{\rho(t_i)} \quad \forall\, t \in T \text{ s.t. } R(t_i, q) \leq k. \tag{3.2.16}$$

Using (3.2.16), (3.2.13) and an argument similar to that used to prove Lemma 3.1.2, we can show (3.2.14) holds for $k \geq 1$. Hence, (3.2.14) holds for $k \geq 0$. Now substituting k for $k+1$ in (3.2.14), we get that, for $k \geq 1$,

$$\tilde{\phi}^{(k)}_{s+1}(t) \in (\tilde{\Phi}_{\rho(t)})_{s+1} \quad \text{for } t = [t_1, \ldots, t_m] \text{ with } R(t_i, q) \leq k-1.$$

Consequently, the t of minimal order for which $\tilde{\phi}^{(k)}_{s+1}(t)$ may fail to be in $(\tilde{\Phi}_{\rho(t)})_{s+1}$ is of the form $t = [t_1]$ with $R(t_1, q) \geq k$, whence $\rho(t_1) \geq q + 2k - 1 = q_k$ by Lemma 2.9, from which it follows that $\rho(t) \geq q_k + 1$. Thus, $\tilde{\phi}^{(k)}_{s+1}(t) \in (\tilde{\Phi}_{\rho(t)})_{s+1}$ for $\rho(t) \leq q_k$. In addition, note that $\tilde{\phi} \in \tilde{\Phi}_\mu$ implies $\tilde{\phi}_{s+1} = 1$ for $\mu \leq p$ by Theorem 2.4. Combining these two observations, we see

$$\tilde{\phi}^{(k)}_{s+1}(t) = 1 \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq \min(p, q_k),$$

which together with Theorem 2.1 implies Case 2 of Theorem 1.1 for $q \geq 0$. For $q = -1$, a similar argument using (3.2.11) instead of (3.2.9) establishes the result. $\square$

Case 2 of Theorem 1.1 can also be sharpened a little by introducing the *R restricted order values*

$$p_k^R = \max\{\mu : \tilde{\phi}_{s+1}(t) = 1 \; \forall \, t = [t_1, \ldots, t_m] \in T \text{ s.t. } \rho(t) \leq \mu \text{ and } R(t_i, 1) \leq k-1 \text{ for } i = 1, \ldots, m\}. \tag{3.2.17}$$

We show below that,

$$y_{n+1}^{(k)} = y(x_{n+1}; x_n, y_n) + O(h^{\min(p_k^R, q_k)+1}) \tag{3.2.18}$$

for $q \geq 0$ and $k \geq 1$, which is a slight improvement over (1.17), since $p_1^R \geq p_2^R \geq \cdots \geq p$, $p_k^R = p$ for $k > p/2$, and, for some RK methods, $p_k^R > p_{k+1}^R$ for some $k \leq p/2$. This result indicates that the order of $y_{n+1}^{(k)}$ may decrease as k increases, as indeed happens in practice in some cases. For $q = -1$, (3.2.18) gives no improvement over (1.17), since the restriction inherent in (3.2.17) does not exclude any trees that are not already excluded in the basic argument used to prove (1.17).

Assume $q \geq 0$ and recall that the assumption of Theorem 1.1 that the initial guess $Y_n^{(0)}$ satisfies (1.9) is equivalent to $\phi^{(0)}(t) = c^{\rho(t)}$ for $\rho(t) \leq q$ (and not just $\phi^{(0)}(t) \in \Omega_{\rho(t)}$ for $\rho(t) \leq q$). To prove (3.2.18) for $k = 1$, consider any $t = [t_1, \ldots, t_m] \in T$ for which $\rho(t_i) \leq q$ for $i = 1, \ldots, m$. By (3.2.2) of Lemma 3.2.1, Theorem 2.4, and the assumptions that $q \geq 0$ and $\phi^{(0)}(t) = c^{\rho(t)}$ for $\rho(t) \leq q$,

$$\tilde{\phi}_{s+1}^{(1)}(t) = \rho(t) b^T \prod_{i=1}^{m} \phi^{(0)}(t_i) = \rho(t) b^T \prod_{i=1}^{m} c^{\rho(t_i)} = \rho(t) b^T c^{\rho(t)-1} = \tilde{\phi}_{s+1}(\hat{t})$$

for $\hat{t} = [\tau^{\rho(t)-1}]$. Hence, $\tilde{\phi}_{s+1}^{(1)}(t) = \tilde{\phi}_{s+1}(\hat{t})$ for $\rho(t) \leq q+1 = q_1$ and $\tilde{\phi}_{s+1}(\hat{t}) = 1$ for $\rho(t) = \rho(\hat{t}) \leq p_1^R$. Therefore, by Theorem 2.1, (3.2.18) holds for $k = 1$. To prove (3.2.18) for $k \geq 2$, first note that, if $q \geq 0$ and $\phi^{(0)}(t) = c^{\rho(t)}$ for $\rho(t) \leq q$, then it follows easily by induction on $k \geq 1$ that, for any $t \in T$ such that $R(t, q) \leq k$, there is an associated $\hat{t} \in T$ such that $\rho(t) = \rho(\hat{t})$, $R(\hat{t}, 1) \leq k$ and $\phi^{(k)}(t) = \phi(\hat{t})$, which of course implies the right side of (3.2.9). Therefore, if $k \geq 2$ and $t = [t_1, \ldots, t_m]$ with $R(t_i, q) \leq k-1$ for $i = 1, \ldots, m$, then by (3.2.2) of Lemma 3.2.1, Theorem 2.4, and the result quoted above

$$\tilde{\phi}_{s+1}^{(k)}(t) = \tilde{\phi}_{s+1}(\hat{t})$$

where $\hat{t} = [\hat{t}_1, \ldots, \hat{t}_m]$ with $\rho(t_i) = \rho(\hat{t}_i)$ and $R(\hat{t}_i, 1) \leq k-1$ for $i = 1, \ldots, m$. Hence, for $k \geq 2$, $\tilde{\phi}_{s+1}^{(k)}(t) = \tilde{\phi}_{s+1}(\hat{t})$ for $\rho(t) \leq q_k$ and $\tilde{\phi}_{s+1}(\hat{t}) = 1$ for $\rho(t) = \rho(\hat{t}) \leq p_k^R$. Therefore, by Theorem 2.1, (3.2.18) holds for $k \geq 2$ also, completing the proof.

It is natural to ask whether the results established in this subsection are tight. It is for this reason that we proved the $\Leftrightarrow$ implications in (3.2.5) of Lemma 3.2.2 and (3.2.8) of Theorem 3.2.3 as well as the converse of (3.2.11): they show that the $\Rightarrow$ implications are tight if A is nonsingular, $c_i \neq 0$ for $i = 1, \ldots, s$ and either $\phi^{(k)}(\emptyset) = \varepsilon_s$ and $\phi^{(k)}(\tau) = c$ or $\phi^{(0)}(\emptyset) = \varepsilon_s$ and $\phi^{(0)}(\tau) = c$, respectively. However, unlike §3.1, it is easy to find examples for which the $\Leftarrow$ directions for (3.2.6), (3.2.7), (3.2.9) and (3.2.10) fail to hold, showing that these implications in the $\Rightarrow$ direction are not tight.

In practice, however, the $\Rightarrow$ implications are often tight even if A is singular or $c_i = 0$ for some $i \in \{1, \ldots, s\}$, although examples can be constructed for which they are not. In particular, if the underlying RK formula is explicit, then, using an argument similar to that employed in §3.1, we can easily show that, if either $q \ge 0$ and $k \ge s+1$ or $q = -1$ and $k \ge s+2$, then $\tilde{\phi}^{(k)}(t) = \tilde{\phi}(t)$ for all $t \in T$. Thus, (3.2.5)–(3.2.7) and (3.2.8)–(3.2.11) are certainly not tight in this case. We also see that, if either $q \ge 0$ and $k \ge s+1$ or $q = -1$ and $k \ge s+2$, then $y_{n+1}^{(k)} = y_{n+1}$, an unusual occurrence for a nonlinear fixed-point iteration.

As in §3.1, it is easy to construct examples for which q is the largest value such that the initial guess $Y_n^{(0)}$ satisfies (1.9), but

$$y_{n+1}^{(k)} = y(x_{n+1}; x_n, y_n) + O(h^{\mu+1}) \tag{3.2.19}$$

for some $k \ge 1$ and some $\mu > \min(p_k^R, q_k) \ge \min(p, q_k)$. However, as noted in §3.1, the B-series for initial guess $Y_n^{(0)}$ often satisfies the stronger assumption (3.1.14). In such cases, the following tighter result holds.

Theorem 3.2.4 *Let the B-series for the initial guess $Y_n^{(0)}$ satisfy (3.1.14) and assume $q \ge 0$ is chosen as large as possible. If either $q \ge 1$ or $q = 0$, $q_k < p_k^R$ and $\zeta(\tau)^2 \ne 1$, then the order of $y_{n+1}^{(k)}$ is exactly $\min(p_k^R, q_k)$, where p_k^R is defined by (3.2.17) and $q_k = q + 2k - 1$.*

Proof: From the discussion following (3.2.17), we know that the order of $y_{n+1}^{(k)}$ is at least $\min(p_k^R, q_k)$. Therefore, we need show only that (3.2.19) does not hold for any $\mu > \min(p_k^R, q_k)$.

First assume $q_k < p_k^R$, and, if $q = 0$, then $\zeta(\tau)^2 \ne 1$. From the hypothesis of the theorem, there is a tree $\hat{t}$ such that $\rho(\hat{t}) = q+1$ and $\zeta(\hat{t}) \ne 1$. If $k = 1$, then by Lemma 3.2.1 and Theorem 2.4

$$\tilde{\phi}_{s+1}^{(1)}([\hat{t}]) = \rho([\hat{t}]) b^T \phi^{(0)}(\hat{t}) = \zeta(\hat{t}) \rho([\hat{t}]) b^T c^{q+1} = \zeta(\hat{t}) \tilde{\phi}_{s+1}([\tau^{q+1}]) = \zeta(\hat{t}) \ne 1,$$

where we have used $\tilde{\phi}_{s+1}([\tau^{q+1}]) = 1$, which follows from $\rho([\tau^{q+1}]) = q+2 = q_1 + 1 \le p_1^R$, $R(\tau, 1) = 0$, and the definition of p_k^R. Since $\tilde{\phi}_{s+1}^{(1)}([\hat{t}]) \ne 1$ for $\rho([\hat{t}]) = q+2 = q_1+1$, the order of $y_{n+1}^{(1)}$ is at most $q_1 = \min(p_1^R, q_1)$, completing the proof for this case. Therefore, assume $k \ge 2$. Using Lemma 3.2.1 and Theorem 2.4 again, one can easily show by induction on $k \ge 1$ that

$$\phi^{(k)}(\hat{t}_{R,k}) = \zeta(\tau)\, \zeta(\hat{t})\, \phi(t_{R,k}), \tag{3.2.20}$$

where $\hat{t}_{R,k}$ and $t_{R,k}$ are defined recursively by

$$\begin{array}{lll} \hat{t}_{R,1} = [\tau, \hat{t}] & \text{and} \quad \hat{t}_{R,k+1} = [\tau, \hat{t}_{R,k}] & \text{for } k \ge 1, \\ t_{R,1} = [\tau^{q+2}] & \text{and} \quad t_{R,k+1} = [\tau, t_{R,k}] & \text{for } k \ge 1. \end{array}$$

Hence, for $k \ge 2$,

$$\begin{aligned} \tilde{\phi}_{s+1}^{(k)}([\hat{t}_{R,k-1}]) &= \rho([\hat{t}_{R,k-1}]) b^T \phi^{(k-1)}(\hat{t}_{R,k-1}) \\ &= \zeta(\tau)\, \zeta(\hat{t})\, \rho([t_{R,k-1}]) b^T \phi(t_{R,k-1}) \\ &= \zeta(\tau)\, \zeta(\hat{t})\, \tilde{\phi}_{s+1}([t_{R,k-1}]). \end{aligned}$$

Now note that $\rho([\hat{t}_{R,k-1}]) = \rho([t_{R,k-1}]) = q_k + 1 \leq p_k^R$ and $R(t_{R,k-1}, 1) = k-1$, from which it follows from the definition of p_k^R that $\tilde{\phi}_{s+1}([t_{R,k-1}]) = 1$, whence

$$\tilde{\phi}_{s+1}^{(k)}([\hat{t}_{R,k-1}]) = \zeta(\tau)\,\zeta(\hat{t}).$$

If $q = 0$, then $\hat{t} = \tau$, since there is one tree only of order 1. Consequently, $\tilde{\phi}_{s+1}^{(k)}([\hat{t}_{R,k-1}]) = (\zeta(\tau))^2 \neq 1$. On the other hand, if $q \geq 1$, then $\zeta(\tau) = 1$ and $\tilde{\phi}_{s+1}^{(k)}([\hat{t}_{R,k-1}]) = \zeta(\hat{t}) \neq 1$, completing the proof for $q_k < p_k^R$ and $q \geq 0$.

Next assume $q \geq 1$. Since we have already proven the result for $q_k < p_k^H$, we need only consider $q \geq 1$ and $p_k^H \leq q_k$. From the definition of p_k^R, there exists a tree $\hat{t} = [\hat{t}_1, \ldots, \hat{t}_m]$ such that $\rho(\hat{t}) = p_k^R + 1$, $R(\hat{t}_i, 1) \leq k-1$ for $i = 1, \ldots, m$ and $\tilde{\phi}_{s+1}(\hat{t}) \neq 1$. Also note that, since $R(\hat{t}_i, 1) \leq k-1$ for $i = 1, \ldots, m$ and $q \geq 1$, we may take $r = 1$ in (3.2.8) to conclude that $\phi^{(k-1)}(\hat{t}_i) = \phi(\hat{t}_i)$ for $i = 1, \ldots, m$. Consequently, by Lemma 3.2.1 and Theorem 2.4,

$$\tilde{\phi}_{s+1}^{(k)}(\hat{t}) = \rho(\hat{t}) b^T \prod_{i=1}^{m} \phi^{(k-1)}(\hat{t}_i) = \rho(\hat{t}) b^T \prod_{i=1}^{m} \phi(\hat{t}_i) = \tilde{\phi}_{s+1}(\hat{t}) \neq 1,$$

whence the order of $y_{n+1}^{(k)}$ is at most $p_k^R = \min(p_k^R, q_k)$. $\square$

3.3 The Full Newton Iteration.

In this subsection, we consider the full Newton iteration, (3.1)–(3.2) with

$$J_n^{(k)} = \operatorname{diag}(f'(Y_{n,1}^{(k)}), \ldots, f'(Y_{n,s}^{(k)})).$$

It follows that the B-series for $\tilde{Y}_n^{(k+1)}$ satisfies

Lemma 3.3.1 *If $k \geq 0$, then $\tilde{\phi}^{(k+1)}(\emptyset) = \varepsilon_{s+1}$. If in addition $\phi^{(k)}(\emptyset) = \varepsilon_s$, then, for $t = [t_1, \ldots, t_m]$ with $\rho(t) \geq 1$,*

$$\begin{aligned} \phi^{(k+1)}(t) &= \rho(t) A \prod_{i=1}^{m} \phi^{(k)}(t_i) + \rho(t) A \sum_{i=1}^{m} \left(\prod_{\substack{j=1 \\ j \neq i}}^{m} \phi^{(k)}(t_j) \right) \left(\phi^{(k+1)}(t_i) - \phi^{(k)}(t_i) \right) && (3.3.1) \\ \tilde{\phi}_{s+1}^{(k+1)}(t) &= \rho(t) b^T \prod_{i=1}^{m} \phi^{(k)}(t_i) && (3.3.2) \end{aligned}$$

where the rightmost $\prod$ in (3.3.1) is taken to be ε_s for $m = 1$. Note in particular that, if $\phi^{(k)}(\emptyset) = \varepsilon_s$, then $\tilde{\phi}^{(k+1)}(\tau) = \tilde{c}$.

Proof: Writing (3.1)–(3.2) as a B-series, we immediately see that $\tilde{\phi}^{(k+1)}(\emptyset) = \varepsilon_{s+1}$ for $k \geq 0$, independently of $\phi^{(k)}$. Therefore, assume $\phi^{(k)}(\emptyset) = \varepsilon_s$ for the remainder of this proof. First note that by Theorem 2.3

$$\begin{aligned} hAJ_n^{(k)}(Y_n^{(k+1)} - Y_n^{(k)}) &= hAf'(B(\phi^{(k)}, y_n))B(\phi^{(k+1)} - \phi^{(k)}, y_n) \\ &= A \sum_{t \in T} \left(\phi^{(k)} \circ (\phi^{(k+1)} - \phi^{(k)}) \right)(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!}. \end{aligned}$$

Using this relation together with Theorem 2.2, we see that (3.1) can be rewritten in B-series notation as

$$\begin{aligned} \sum_{t \in T} \phi^{(k+1)}(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} &= y_n \varepsilon_s + A \sum_{t \in T} \phi'^{(k)}(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} \\ &+ A \sum_{t \in T} \left(\phi^{(k)} \circ (\phi^{(k+1)} - \phi^{(k)}) \right)(t) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!} \end{aligned} \qquad (3.3.3)$$

from which (3.3.1) follows by comparing corresponding terms in (3.3.3) and using Theorems 2.2 and 2.3 once more. Also, (3.3.2) follows from (3.2) by an argument similar to that used in §3.2 to prove (3.2.2) of Lemma 3.2.1. Finally, since $\tau = [\emptyset]$ and $\phi^{(k)}(\emptyset) = \phi^{(k+1)}(\emptyset) = \varepsilon_s$, it follows from (3.3.1), (3.3.2) and Theorem 2.4 that

$$\begin{aligned} \phi^{(k+1)}(\tau) &= A\phi^{(k+1)}(\emptyset) = A\varepsilon_s = c, \\ \tilde{\phi}_{s+1}^{(k+1)}(\tau) &= b^T \phi^{(k)}(\emptyset) = b^T \varepsilon_s = \tilde{c}_{s+1}, \end{aligned}$$

whence $\tilde{\phi}^{(k+1)}(\tau) = \tilde{c}$. □

Lemma 3.3.2 *For r, q, k, $l \geq 0$,*

$$\begin{aligned} &\phi^{(k)}(t) = \phi(t)\ \forall\, t \in T\ s.t.\ D(t,r,r) \leq l \\ &\qquad \Rightarrow \phi^{(k+1)}(t) = \phi(t)\ \forall\, t \in T\ s.t.\ D(t,r,r) \leq l+1, \end{aligned} \tag{3.3.4}$$

$$\begin{aligned} &\begin{cases} \phi^{(k)}(t) = \phi(t)\ \forall\, t \in T\ s.t.\ D(t,r,r) \leq l \quad and \\ \phi^{(k)}(t) \in \Upsilon_{\rho(t)}\ \forall\, t \in T\ s.t.\ D(t,r,q) \leq l \end{cases} \\ &\qquad \Rightarrow \phi^{(k+1)}(t) \in \Psi_{\rho(t)}\ \forall\, t \in T\ s.t.\ D(t,r,q) \leq l+1. \end{aligned} \tag{3.3.5}$$

Moreover, if $\phi^{(k)}(\emptyset) = \varepsilon_s$ and A is nonsingular, then the $\Rightarrow$ in (3.3.4) can be replaced by a $\Leftrightarrow$.

Proof: We first prove (3.3.4) in the $\Rightarrow$ direction by induction on $\rho(t) \geq 0$. To begin, note that, if $\rho(t) = 0$, then $t = \emptyset$ and $\phi^{(k+1)}(\emptyset) = \varepsilon_s = \phi(\emptyset)$ by Lemma 3.3.1 and Theorem 2.4, whence the conclusion of (3.3.4) holds for $\rho(t) = 0$. To show this implication holds for $\rho(t) \geq 1$ also, first note that $\phi^{(k)}(\emptyset) = \phi(\emptyset)$ from the hypothesis of (3.3.4), since $D(\emptyset, r, r) \leq l$ for r, $l \geq 0$ by Definition 2.5. In addition, since $\phi(\emptyset) = \varepsilon_s$ by Theorem 2.4, $\phi^{(k)}(\emptyset) = \varepsilon_s$, whence (3.3.1) of Lemma 3.3.1 holds. Let $\bar{\rho} \geq 1$ and assume by the induction hypothesis that the conclusion of (3.3.4) holds for any tree t satisfying $D(t,r,r) \leq l+1$ and $\rho(t) < \bar{\rho}$. We must show that $\phi^{(k+1)}(\bar{t}) = \phi(\bar{t})$ for any tree $\bar{t}$ satisfying $D(\bar{t},r,r) \leq l+1$ and $\rho(\bar{t}) = \bar{\rho} \geq 1$. Since $\rho(\bar{t}) \geq 1$, $\bar{t} = [t_1, \ldots, t_m]$ for some t_i satisfying $\rho(t_i) < \bar{\rho}$ for $i = 1, \ldots, m$. Moreover, by Lemma 2.8, we may renumber the t_i so that $D(t_i, r, r) \leq l$ for $i = 1, \ldots, m-1$ and $Dt_m, r, r) \leq l+1$. Consequently, $\phi^{(k)}(t_i) = \phi(t_i)$ for $i = 1, \ldots, m-1$ by the hypothesis of (3.3.4) and $\phi^{(k+1)}(t_i) = \phi(t_i)$ for $i = 1, \ldots, m$ by the induction hypothesis. Thus, by (3.3.1) of Lemma 3.3.1 and Theorem 2.4,

$$\begin{aligned} \phi^{(k+1)}(\bar{t}) &= \rho(\bar{t}) A \prod_{i=1}^{m} \phi^{(k)}(t_i) + \rho(\bar{t}) A \sum_{i=1}^{m} \left(\prod_{\substack{j=1 \\ j \neq i}}^{m} \phi^{(k)}(t_j) \right) \left(\phi^{(k+1)}(t_i) - \phi^{(k)}(t_i) \right) \\ &= \rho(\bar{t}) A \left(\prod_{i=1}^{m-1} \phi(t_i) \right) \phi^{(k)}(t_m) + \rho(\bar{t}) A \left(\prod_{j=1}^{m-1} \phi(t_j) \right) \left(\phi(t_m) - \phi^{(k)}(t_m) \right) \\ &= \rho(\bar{t}) A \prod_{i=1}^{m} \phi(t_i) \\ &= \phi(\bar{t}) \end{aligned}$$

completing the induction proof of (3.3.4) in the $\Rightarrow$ direction. The proof of (3.3.5) is similar, although a little more complex.

Assume $\phi^{(k)}(\emptyset) = \varepsilon_s$ and A is nonsingular. To prove (3.3.4) in the $\Leftarrow$ direction, first note that $\phi^{(k)}(\emptyset) = \phi(\emptyset)$, since $\phi(\emptyset) = \varepsilon_s$ too by Theorem 2.4. Therefore (3.3.4) in the $\Leftarrow$ direction holds for $t = \emptyset$. To see that it also holds for any t satisfying $\rho(t) \geq 1$ and $D(t,r,r) \leq l$, let $\bar{t} = [t,t]$ for any such t. Since $D(\bar{t},r,r) \leq l+1$ by Definition 2.5, $\phi^{(k+1)}(\bar{t}) = \phi(\bar{t})$ by the right side of (3.3.4). Now note that, by (3.3.1) of Lemma 3.3.1,

$$\phi^{(k+1)}(\bar{t}) = \rho(\bar{t}) A \left(\phi^{(k)}(t) \right)^2 + 2\rho(\bar{t}) A \left(\phi^{(k)}(t) \left(\phi^{(k+1)}(t) - \phi^{(k)}(t) \right) \right)$$

$$
\begin{aligned}
&= \rho(\bar{t})A\left(2\phi^{(k)}(t)\phi^{(k+1)}(t) - \left(\phi^{(k)}(t)\right)^2\right) \\
&= \rho(\bar{t})A\left(2\phi^{(k)}(t)\phi(t) - \left(\phi^{(k)}(t)\right)^2\right)
\end{aligned}
$$

and by Theorem 2.4

$$\phi(\bar{t}) = \rho(\bar{t})A\left(\phi(t)\right)^2.$$

Therefore, since $\phi^{(k+1)}(\bar{t}) = \phi(\bar{t})$, $\rho(\bar{t}) > 0$ and A is nonsingular,

$$\left(\phi(t)\right)^2 = 2\phi^{(k)}(t)\phi(t) - \left(\phi^{(k)}(t)\right)^2$$

which can be rewritten as

$$0 = \left(\phi(t)\right)^2 - 2\phi^{(k)}(t)\phi(t) + \left(\phi^{(k)}(t)\right)^2 = \left(\phi(t) - \phi^{(k)}(t)\right)^2,$$

whence $\phi^{(k)}(t) = \phi(t)$, as required. $\square$

Theorem 3.3.3 *If r, $q \geq 0$ and $k \geq 1$, then*

$$
\begin{aligned}
&\phi^{(0)}(t) = \phi(t)\ \forall\, t \in T \textit{ s.t. } \rho(t) \leq r \\
&\qquad \Rightarrow \phi^{(k)}(t) = \phi(t)\ \forall\, t \in T \textit{ s.t. } D(t, r, r) \leq k,
\end{aligned}
\tag{3.3.6}
$$

$$
\begin{aligned}
&\begin{cases} \phi^{(0)}(t) = \phi(t)\ \forall\, t \in T \textit{ s.t. } \rho(t) \leq r \textit{ and} \\ \phi^{(0)}(t) \in \Upsilon_{\rho(t)}\ \forall\, t \in T \textit{ s.t. } \rho(t) \leq q \end{cases} \\
&\qquad \Rightarrow \phi^{(k)}(t) \in \Psi_{\rho(t)}\ \forall\, t \in T \textit{ s.t. } D(t, r, q) \leq k.
\end{aligned}
\tag{3.3.7}
$$

For the special case $r\,(= q) = -1$ and $k \geq 1$,

$$\phi^{(k)}(t) = \phi(t) \quad \forall\, t \in T \textit{ s.t. } D(t) \leq k - 1. \tag{3.3.8}$$

Moreover, if $\phi^{(0)}(\emptyset) = \varepsilon_s$ and A is nonsingular, then the $\Rightarrow$ in (3.3.6) can be replaced by a $\Leftrightarrow$. In addition, if $\phi^{(1)}(\tau) \neq c$, and A is nonsingular, then (3.3.8) is tight in the sense that $\phi^{(k)}(t_{D,k}) \neq \phi(t_{D,k})$ for $k \geq 1$, where $t_{D,k}$ is defined recursively by $t_{D,1} = \tau$ and $t_{D,k+1} = [t_{D,k}, t_{D,k}]$ for $k \geq 1$, whence $D(t_{D,k}) = k$.

Proof: Since $\rho(t) \leq \mu$ if and only if $D(t, \mu, \mu) = 0$ and $\rho(t) \leq \nu$ if and only if $D(t, \mu, \nu) = 0$, (3.3.6) and (3.3.7) follow directly by induction on k from Lemma 3.3.2 with $l = k$. For $r = -1$, $\tilde{\phi}^{(1)}(\emptyset) = \varepsilon_{s+1} = \tilde{\phi}(\emptyset)$ by Lemma 3.3.1 and Theorem 2.4. Therefore (3.3.8) follows by induction on k from Lemma 3.3.2 with $l = k - 1$.

We use induction on $k \geq 1$ to show $\phi^{(k)}(t_{D,k}) \neq \phi(t_{D,k})$ if $\phi^{(1)}(\tau) \neq c$ and A is nonsingular. To establish the base case $k = 1$, note that $t_{D,1} = \tau$ by definition and $\phi(\tau) = c$ by Theorem 2.4. Therefore, $\phi^{(1)}(t_{D,1}) \neq \phi(t_{D,1})$ since $\phi^{(1)}(\tau) \neq c$. For the induction step, assume $k \geq 2$ and $\phi^{(k-1)}(t_{D,k-1}) \neq \phi(t_{D,k-1})$. However, $\phi^{(k)}(t_{D,k-1}) = \phi(t_{D,k-1})$ by (3.3.8) since $D(t_{D,k-1}) = k - 1$. It follows from these relations, Lemma 3.3.1 and Theorem 2.4 that

$$
\begin{aligned}
\phi(t_{D,k}) - \phi^{(k)}(t_{D,k}) &= \rho(t_{D,k})A\left(\phi(t_{D,k-1})\right)^2 - \rho(t_{D,k})A\left(\phi^{(k-1)}(t_{D,k-1})\right)^2 \\
&\qquad -2\rho(t_{D,k})A\phi^{(k-1)}(t_{D,k-1})\left(\phi^{(k)}(t_{D,k-1}) - \phi^{(k-1)}(t_{D,k-1})\right) \\
&= \rho(t_{D,k})A\left[\left(\phi(t_{D,k-1})\right)^2 - 2\phi^{(k-1)}(t_{D,k-1})\phi(t_{D,k-1}) + \left(\phi^{(k-1)}(t_{D,k-1})\right)^2\right] \\
&= \rho(t_{D,k})A\left[\phi(t_{D,k-1}) - \phi^{(k-1)}(t_{D,k-1})\right]^2 \neq 0
\end{aligned}
$$

where for the last inequality we have used the induction hypothesis $\phi^{(k-1)}(t_{D,k-1}) \neq \phi(t_{D,k-1})$ together with $\rho(t_{D,k}) > 0$ and A nonsingular. $\square$

Proof of Case 3 of Theorem 1.1. Assume first that $q \geq 0$. As noted at the start of §3, the assumption of Theorem 1.1 that the initial guess $Y_n^{(0)}$ satisfies (1.9) is equivalent to (3.3). That is, we may assume

$$\phi^{(0)}(t) = c^{\rho(t)} \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq q. \tag{3.3.9}$$

In particular, $\phi^{(0)}(\emptyset) = \varepsilon_s$. In addition, Lemma 3.3.1 ensures that $\tilde{\phi}^{(k)}(\emptyset) = \varepsilon_{s+1}$ for $k \geq 1$, whence $\tilde{\phi}^{(k)}_{s+1}(\emptyset) = 1$ for $k \geq 1$. Moreover, it follows from Lemma 3.3.1 that

$$\tilde{\phi}^{(k+1)}_{s+1}(t) = \rho(t) b^T \prod_{i=1}^{m} \phi^{(k)}(t_i) \tag{3.3.10}$$

for $t = [t_1, \ldots, t_m]$ and $k \geq 0$. We will now show that, for $k \geq 0$,

$$\tilde{\phi}^{(k+1)}_{s+1}(t) \in (\tilde{\Psi}_{\rho(t)})_{s+1} \quad \text{for } t = [t_1, \ldots, t_m] \text{ with } D(t_i, r, q) \leq k. \tag{3.3.11}$$

To this end, first note that (3.3.9) and Definition 2.5 ensure that

$$\phi^{(0)}(t_i) = c^{\rho(t_i)} \quad \forall\, t \in T \text{ s.t. } D(t_i, r, q) = 0. \tag{3.3.12}$$

Using (3.3.12), (3.3.10) for $k = 0$, and an argument similar to that used to prove Lemma 3.1.2, we can show (3.3.11) holds for $k = 0$. To prove (3.3.11) for $k \geq 1$, first note that by Theorem 2.5

$$\phi(t) = c^{\rho(t)} \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq \eta,$$

where η is the stage order (defined by (1.6) or (2.4)) of the underlying RK formula. This together with (3.3.9) ensures that

$$\phi^{(0)}(t) = \phi(t) \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq r = \min(q, \eta).$$

Moreover, since $c^{\rho(t)} \in \Upsilon_{\rho(t)}$ for all $t \in T$,

$$\phi^{(0)}(t) \in \Upsilon_{\rho(t)} \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq q.$$

Therefore, by (3.3.7),

$$\phi^{(k)}(t) \in \Psi_{\rho(t)} \quad \forall\, t \in T \text{ s.t. } D(t, r, q) \leq k. \tag{3.3.13}$$

Using (3.3.13), (3.3.10) and an argument similar to that used to prove Lemma 3.1.2, we can show (3.3.11) holds for $k \geq 1$. Hence, (3.3.11) holds for $k \geq 0$. Now substituting k for $k+1$ in (3.3.11), we get that, for $k \geq 1$,

$$\tilde{\phi}^{(k)}_{s+1}(t) \in (\tilde{\Psi}_{\rho(t)})_{s+1} \quad \text{for } t = [t_1, \ldots, t_m] \text{ with } D(t_i, r, q) \leq k - 1.$$

Consequently, the t of minimal order for which $\tilde{\phi}^{(k)}_{s+1}(t)$ may fail to be in $(\tilde{\Psi}_{\rho(t)})_{s+1}$ is of the form $t = [t_1]$ with $D(t_1, r, q) \geq k$, whence $\rho(t_1) \geq 2^{k-1}(r+2) + q - r - 1 = q_k$ by Lemma 2.9, from which it follows that $\rho(t) \geq q_k + 1$. Thus, $\tilde{\phi}^{(k)}_{s+1}(t) \in (\tilde{\Psi}_{\rho(t)})_{s+1}$ for $\rho(t) \leq q_k$. In addition, note that $\tilde{\phi} \in \tilde{\Psi}_\mu$ implies $\tilde{\phi}_{s+1} = 1$ for $\mu \leq p$ by Lemma 2.6. Combining these two observations, we see

$$\tilde{\phi}^{(k)}_{s+1}(t) = 1 \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq \min(p, q_k),$$

which together with Theorem 2.1 implies Case 3 of Theorem 1.1 for $q \geq 0$. For $q = -1$, a similar argument using (3.3.8) instead of (3.3.7) establishes the result. $\square$

Case 3 of Theorem 1.1 can also be sharpened a little by introducing the D *restricted order values*

$$\begin{aligned} p^D_k = \max\{\mu \,:\, & \tilde{\phi}_{s+1}(t) = 1 \;\forall\, t = [t_1, \ldots, t_m] \in T \text{ s.t. } \rho(t) \leq \mu \text{ and} \\ & D(t_i, 1, 1) \leq k - 1 \text{ for } i = 1, \ldots, m\}. \end{aligned} \tag{3.3.14}$$

It is easy to show by induction on k that we can extend (3.3.13) to

$$\phi^{(k)}(t) \in \Psi^{(k)}_{\rho(t)} \quad \forall\, t \in T \text{ s.t. } D(t, r, q) \leq k \tag{3.3.15}$$

where

$$\begin{aligned} \Psi^{(k)}_r \;=\; & \{\phi(t) + \sum_{j=1}^{\nu} \kappa_j \left(\phi(t_{j,1}) - \phi(t_{j,2})\right) : \rho(t) = \rho(t_{j,1}) = \rho(t_{j,2}) = r,\ j = 1, \ldots, \nu; \\ & D(t, 1, 1) \leq k,\ D(t_{j,1}, 1, 1) \leq k,\ D(t_{j,2}, 1, 1) \leq k,\ j = 1, \ldots, \nu; \\ & \kappa_j \in \mathbb{R},\ j = 1, \ldots, \nu\}. \end{aligned}$$

Using (3.3.15) in place of (3.3.13), we can extend the proof given above of Case 3 of Theorem 1.1 to show that

$$y^{(k)}_{n+1} = y(x_{n+1}; x_n, y_n) + O(h^{\min(p^D_k, q_k)+1}) \tag{3.3.16}$$

for $q \geq 0$ and $k \geq 1$. The error estimate (3.3.16) is a slight improvement over (1.17), since $p^D_1 \geq p^D_2 \geq \cdots \geq p$ and, for some RK formulas, $p^D_k > p^D_{k+1}$ for some $k \geq 1$. (However, $p^D_k = p$ for $k > 1 + \log_2 \frac{p+2}{3}$.) Equation (3.3.16) suggests that the order of $y^{(k)}_{n+1}$ may decrease as k increases, as indeed happens in practice in some cases. For $q = -1$, (3.3.16) gives no improvement over (1.17), since the restriction inherent in (3.3.14) does not exclude any trees that are not already excluded in the basic argument used above to prove (1.17).

It is natural to ask whether the results established in this subsection are tight. It is for this reason that we proved the $\Leftrightarrow$ implications in (3.3.4) of Lemma 3.3.2 and (3.3.6) of Theorem 3.3.3 as well as the converse of (3.3.8): they show that the $\Rightarrow$ implications are tight if A is nonsingular and either $\phi^{(k)}(\emptyset) = \varepsilon_s$ or $\phi^{(0)}(\emptyset) = \varepsilon_s$, respectively. However, as in §3.2, it is easy to find examples for which the $\Leftarrow$ directions for (3.3.5) and (3.3.7) fail to hold, showing that these implications in the $\Rightarrow$ direction are not tight.

In practice, however, the $\Rightarrow$ implications are often tight even if A is singular, although examples can be constructed for which they are not. In particular, if the underlying RK formula is explicit, then, using an argument similar to that employed in §3.1, we can easily show that, if either $q \geq 0$ and $k \geq s+1$ or $q = -1$ and $k \geq s+2$, then $\tilde{\phi}^{(k)}(t) = \tilde{\phi}(t)$ for all $t \in T$. Thus, (3.3.4)–(3.3.5) and (3.3.6)–(3.3.8) are certainly not tight in this case. We also see that, if either $q \geq 0$ and $k \geq s+1$ or $q = -1$ and $k \geq s+2$, then $y_{n+1}^{(k)} = y_{n+1}$, an unusual occurrence for a nonlinear fixed-point iteration.

As in §3.1, it is easy to construct examples for which q is the largest value such that the initial guess $Y_n^{(0)}$ satisfies (1.9), but

$$y_{n+1}^{(k)} = y(x_{n+1}; x_n, y_n) + O(h^{\mu+1}) \tag{3.3.17}$$

for some $k \geq 1$ and some $\mu > \min(p_k^D, q_k) \geq \min(p, q_k)$. However, as noted in §3.1, the B-series for initial guess $Y_n^{(0)}$ often satisfies the stronger assumption (3.1.14). In this case, the following theorem holds. Note that, unlike Case 3 of Theorem 1.1, it shows that we benefit from a high-order predictor even if the formula's stage order is low. Examples 4.3 and 4.4 in the next section illustrate this point. Thus, Theorem 3.3.4 is a much larger improvement over Case 3 of Theorem 1.1 than either Theorem 3.1.4 or Theorem 3.2.4 is over Case 1 or Case 2, respectively, of Theorem 1.1.

Theorem 3.3.4 *Let the B-series for the initial guess $Y_n^{(0)}$ satisfy (3.1.14) and assume that the associated q is chosen as large as possible and that $q \geq 0$. Then, for $k \geq 1$, the order of $y_{n+1}^{(k)}$ is at least $\min(p_k^D, \bar{q}_k)$, where p_k^D is defined by (3.3.14) and $\bar{q}_k = 2^{k-1}(q+2)-1$. If either $q \geq 1$ or $\bar{q}_k < p_k^D$, then the order of $y_{n+1}^{(k)}$ is exactly $\min(p_k^D, \bar{q}_k)$.*

Before proving this result, we establish the following lemma.

Lemma 3.3.5 *Assume (3.1.14) holds, the associated q is chosen as large as possible and $q \geq 0$.*

(a) *For any $t \in T$ and any $k \geq 1$,*

$$\phi^{(k)}(t) = \sum_{\nu=1}^{\mathcal{N}_{t,k}} \kappa_\nu^{(k)}(t)\phi(t_\nu^{(k)}), \tag{3.3.18}$$

where each $\kappa_\nu^{(k)}(t) : T \to \mathbb{R}$ and each $t_\nu^{(k)}$ is a tree satisfying $\rho(t_\nu^{(k)}) = \rho(t)$ and $D(t_\nu^{(k)}, 1, 1) \leq k$.

(b) *If $k \geq 1$, then*

$$\mathcal{K}^{(k)}(t) \equiv \sum_{\nu=1}^{\mathcal{N}_{t,k}} \kappa_\nu^{(k)}(t) = 1 \quad \forall\, t \in T \text{ s.t. } D(t,q,q) \leq k. \tag{3.3.19}$$

(c) *Let $\hat{t}$ be any tree for which $\rho(\hat{t}) = q+1$ and $\zeta(\hat{t}) \neq 1$. For $k \geq 1$, define $t_{D,k}$ recursively by*

$$t_{D,1} = [\hat{t}, \hat{t}] \quad \text{and} \quad t_{D,k+1} = [t_{D,k}, t_{D,k}] \quad \text{for } k \geq 1.$$

Then

$$\mathcal{K}^{(k)}(t_{D,k}) \equiv \sum_{\nu=1}^{\mathcal{N}_{t,k}} \kappa_\nu^{(k)}(t_{D,k}) = 1 - (1-\zeta(\hat{t}))^{2^k} \neq 1. \tag{3.3.20}$$

Proof: We prove part (a) by an inner and an outer induction: the outer induction is on k while the inner induction, used to prove both the base case ($k = 1$) and the induction step ($k \geq 2$) of the outer induction, is on $\rho(t)$.

To prove (3.3.18) for $k = 1$, assume first that $\rho(t) = 0$. Therefore, $t = \emptyset$ and $\phi^{(1)}(\emptyset) = \varepsilon_s = \phi(\emptyset)$ by Lemma 3.3.1 and Theorem 2.4. Therefore (3.3.18) holds for $k = 1$ and $\rho(t) = 0$.

Next assume (3.3.18) holds for $k = 1$ and $\rho(t) < \bar{\rho}$ for some $\bar{\rho} \geq 1$. Let $t = [t_1, \ldots, t_m]$ be any tree satisfying $\rho(t) = \bar{\rho}$. Since $\rho(t_i) < \bar{\rho}$ for $i = 1, \ldots, m$, it follows from the inner induction hypothesis that

$$\phi^{(1)}(t_i) = \sum_{\nu=1}^{\mathcal{N}_{t_i,1}} \kappa_\nu^{(1)}(t_i)\phi(t_{i,\nu}^{(1)}), \tag{3.3.21}$$

where $\rho(t_i) = \rho(t_{i,\nu}^{(1)})$ and $D(t_{i,\nu}^{(1)}, 1, 1) \leq 1$ for $\nu = 1, \ldots, \mathcal{N}_{t_i,1}$ and $i = 1, \ldots, m$. Since (3.1.14) holds with $q \geq 0$, $\phi^{(0)}(\emptyset) = \varepsilon_s$, whence by Lemma 3.3.1

$$\begin{aligned}
\phi^{(1)}(t) &= \rho(t)A\prod_{i=1}^{m}\phi^{(0)}(t_i) + \rho(t)A\sum_{i=1}^{m}\left(\prod_{\substack{j=1\\j\neq i}}^{m}\phi^{(0)}(t_j)\right)\left(\phi^{(1)}(t_i) - \phi^{(0)}(t_i)\right) \\
&= (1-m)\rho(t)A\prod_{i=1}^{m}\phi^{(0)}(t_i) + \rho(t)A\sum_{i=1}^{m}\left(\prod_{\substack{j=1\\j\neq i}}^{m}\phi^{(0)}(t_j)\right)\phi^{(1)}(t_i) \\
&= (1-m)\rho(t)A\prod_{i=1}^{m}\zeta(t_i)c^{\rho(t_i)} + \rho(t)A\sum_{i=1}^{m}\left(\prod_{\substack{j=1\\j\neq i}}^{m}\zeta(t_j)c^{\rho(t_j)}\right)\phi^{(1)}(t_i). \qquad (3.3.22)
\end{aligned}$$

Letting

$$\bar{\kappa} = \prod_{i=1}^{m}\zeta(t_i) \quad \text{and} \quad \bar{\kappa}_i = \prod_{\substack{j=1\\j\neq i}}^{m}\zeta(t_j)$$

noting $\sum_{i=1}^{m}\rho(t_i) = \rho(t) - 1$ and substituting the right side of (3.3.21) for $\phi^{(1)}(t_i)$ into (3.3.22), we get

$$\begin{aligned}
\phi^{(1)}(t) &= (1-m)\bar{\kappa}\rho(t)Ac^{\rho(t)-1} + \rho(t)A\sum_{i=1}^{m}\bar{\kappa}_i c^{\rho(t)-\rho(t_i)-1}\left(\sum_{\nu=1}^{\mathcal{N}_{t_i,1}}\kappa_\nu^{(1)}(t_i)\phi(t_{i,\nu}^{(1)})\right) \\
&= (1-m)\bar{\kappa}\rho(t)Ac^{\rho(t)-1} + \sum_{i=1}^{m}\sum_{\nu=1}^{\mathcal{N}_{t_i,1}}\left(\bar{\kappa}_i\kappa_\nu^{(1)}(t_i)\right)\left(\rho(t)Ac^{\rho(t)-\rho(t_i)-1}\phi(t_{i,\nu}^{(1)})\right) \\
&= \kappa_{0,1}^{(1)}(t)\phi([\tau^{\rho(t)-1}]) + \sum_{i=1}^{m}\sum_{\nu=1}^{\mathcal{N}_{t_i,1}}\kappa_{i,\nu}^{(1)}(t)\phi([\tau^{\rho(t)-\rho(t_i)-1}, t_{i,\nu}^{(1)}]) \qquad (3.3.23)
\end{aligned}$$

where $\kappa_{0,1}^{(1)}(t) = (1-m)\bar{\kappa}$ and $\kappa_{i,\nu}^{(1)}(t) = \bar{\kappa}_i\kappa_\nu^{(1)}(t_i)$. Now note that (3.3.23) is of the form (3.3.18), each $\kappa_{i,\nu}^{(1)}(t) : T \to \mathbb{R}$, $\rho([\tau^{\rho(t)-1}]) = \rho(t)$, $D([\tau^{\rho(t)-1}], 1, 1) \leq 1$, $\rho([\tau^{\rho(t)-\rho(t_i)-1}, t_{i,\nu}^{(1)}]) =$

$\rho(t)$ since $\rho(t_{i,\nu}^{(1)}) = \rho(t_i)$ and $D([\tau^{\rho(t)-\rho(t_i)-1}, t_{i,\nu}^{(1)}], 1, 1) \leq 1$ since $D(t_{i,\nu}^{(1)}, 1, 1) \leq 1$ and $D(\tau, 1, 1) = 0$. Therefore (3.3.18) holds for $k = 1$.

For the outer induction step, assume that (3.3.18) holds for $k-1$ where $k \geq 2$. We use an inner induction on $\rho(t)$ to show (3.3.18) holds for k.

If $\rho(t) = 0$, then $t = \emptyset$ and $\phi^{(k)}(\emptyset) = \varepsilon_s = \phi(\emptyset)$ by Lemma 3.3.1 and Theorem 2.4. Therefore (3.3.18) holds for k and $\rho(t) = 0$.

For the inner induction step, assume also that (3.3.18) holds for k and $\rho(t) < \bar{\rho}$ for some $\bar{\rho} \geq 1$. To complete the proof, we will show that (3.3.18) holds for any $t = [t_1, \ldots, t_m] \in T$ satisfying $\rho(t) = \bar{\rho}$.

By our outer induction hypothesis

$$\phi^{(k-1)}(t_i) = \sum_{\nu=1}^{\mathcal{N}_{t_i,k-1}} \kappa_\nu^{(k-1)}(t_i)\phi(t_{i,\nu}^{(k-1)})$$

where $\rho(t_i) = \rho(t_{i,\nu}^{(k-1)})$ and $D(t_{i,\nu}^{(k-1)}, 1, 1) \leq k-1$ for $\nu = 1, \ldots, \mathcal{N}_{t_i,k-1}$ and $i = 1, \ldots, m$. Similarly, since $\rho(t_i) < \bar{\rho}$, by our inner induction hypothesis

$$\phi^{(k)}(t_i) = \sum_{\nu=1}^{\mathcal{N}_{t_i,k}} \kappa_\nu^{(k)}(t_i)\phi(t_{i,\nu}^{(k)})$$

where $\rho(t_i) = \rho(t_{i,\nu}^{(k)})$ and $D(t_{i,\nu}^{(k)}, 1, 1) \leq k$ for $\nu = 1, \ldots, \mathcal{N}_{t_i,k}$ and $i = 1, \ldots, m$. Since $k - 1 \geq 1$ and (3.1.14) holds with $q \geq 0$, $\phi^{(k-1)}(\emptyset) = \varepsilon_s$, whence by Lemma 3.3.1

$$\begin{aligned}
\phi^{(k)}(t) &= \rho(t)A\prod_{i=1}^m \phi^{(k-1)}(t_i) + \rho(t)A\sum_{i=1}^m \left(\prod_{\substack{j=1\\j\neq i}}^m \phi^{(k-1)}(t_j)\right)\left(\phi^{(k)}(t_i) - \phi^{(k-1)}(t_i)\right) \\
&= (1-m)\rho(t)A\prod_{i=1}^m \phi^{(k-1)}(t_i) + \rho(t)A\sum_{i=1}^m \left(\prod_{\substack{j=1\\j\neq i}}^m \phi^{(k-1)}(t_j)\right)\phi^{(k)}(t_i) \\
&= (1-m)\rho(t)A\prod_{i=1}^m \left[\sum_{\nu=1}^{\mathcal{N}_{t_i,k-1}} \kappa_\nu^{(k-1)}(t_i)\phi(t_{i,\nu}^{(k-1)})\right] \qquad (3.3.24) \\
&\quad +\rho(t)A\sum_{i=1}^m \left(\prod_{\substack{j=1\\j\neq i}}^m \left[\sum_{\nu=1}^{\mathcal{N}_{t_j,k-1}} \kappa_\nu^{(k-1)}(t_j)\phi(t_{j,\nu}^{(k-1)})\right]\right)\left[\sum_{\nu=1}^{\mathcal{N}_{t_i,k}} \kappa_\nu^{(k)}(t_i)\phi(t_{i,\nu}^{(k)})\right] \qquad (3.3.25)
\end{aligned}$$

Each term in (3.3.24) is of the form

$$\begin{aligned}
(1-m)\rho(t)A\prod_{i=1}^m \kappa_{\nu_{i,k-1}}^{(k-1)}(t_i)\phi(t_{i,\nu_{i,k-1}}^{(k-1)}) &= \left((1-m)\prod_{i=1}^m \kappa_{\nu_{i,k-1}}^{(k-1)}(t_i)\right)\left(\rho(t)A\prod_{i=1}^m \phi(t_{i,\nu_{i,k-1}}^{(k-1)})\right) \\
&= \kappa_\nu^{(k)}(t)\phi([t_{1,\nu_{1,k-1}}^{(k-1)}, \ldots, t_{m,\nu_{m,k-1}}^{(k-1)}])
\end{aligned}$$

where each $\nu_{i,k-1} \in \{1, \ldots, \mathcal{N}_{t_i,k-1}\}$ and

$$\kappa_\nu^{(k)}(t) = (1-m) \prod_{i=1}^m \kappa_{\nu_{i,k-1}}^{(k-1)}(t_i).$$

Note that $\rho([t_{1,\nu_{1,k-1}}^{(k-1)}, \ldots, t_{m,\nu_{m,k-1}}^{(k-1)}]) = \rho(t)$ since $t = [t_1, \ldots, t_m]$ and by the outer induction hypothesis $\rho(t_{i,\nu_{i,k-1}}^{(k-1)}) = \rho(t_i)$ for $i = 1, \ldots, m$. Also, $D([t_{1,\nu_{1,k-1}}^{(k-1)}, \ldots, t_{m,\nu_{m,k-1}}^{(k-1)}], 1, 1) \leq k$ since by the outer induction hypothesis $D(t_{i,\nu_{i,k-1}}^{(k-1)}, 1, 1) \leq k-1$ for $i = 1, \ldots, m$. Thus, each term in (3.3.24) is of the form required for (3.3.18).

Similarly each term in (3.3.25) is of the form

$$\begin{aligned} \rho(t) A \prod_{i=1}^m \kappa_{\nu_{i,k_i}}^{(k_i)}(t_i) \phi(t_{i,\nu_i}^{(k_i)}) &= \left(\prod_{i=1}^m \kappa_{\nu_{i,k_i}}^{(k_i)}(t_i)\right) \left(\rho(t) A \prod_{i=1}^m \phi(t_{i,\nu_{i,k_i}}^{(k_i)})\right) \\ &= \kappa_\nu^{(k)}(t) \phi([t_{1,\nu_{1,k_1}}^{(k_1)}, \ldots, t_{m,\nu_{m,k_m}}^{(k_m)}]) \end{aligned}$$

where, for $i, j \in \{1, \ldots, m\}$, exactly one $k_i = k$, all other $k_j = k-1$, each $\nu_{i,k-1} \in \{1, \ldots, \mathcal{N}_{t_i,k-1}\}$, each $\nu_{i,k} \in \{1, \ldots, \mathcal{N}_{t_i,k}\}$ and

$$\kappa_\nu^{(k)}(t) = \prod_{i=1}^m \kappa_{\nu_{i,k_i}}^{(k_i)}(t_i).$$

Note that $\rho([t_{1,\nu_{1,k_1}}^{(k_1)}, \ldots, t_{m,\nu_{m,k_m}}^{(k_m)}]) = \rho(t)$ since $t = [t_1, \ldots, t_m]$ and by the induction hypotheses $\rho(t_{i,\nu_{i,k_i}}^{(k_i)}) = \rho(t_i)$ for $i = 1, \ldots, m$ independently of whether $k_i = k$ or $k-1$. Also, $D([t_{1,\nu_{1,k_1}}^{(k_1)}, \ldots, t_{m,\nu_{m,k_m}}^{(k_m)}], 1, 1) \leq k$ since, as noted above, exactly one $k_i = k$ and all other $k_j = k-1$ and by the inner induction hypothesis $D(t_{i,\nu_{i,k}}^{(k)}, 1, 1) \leq k$ for $i = 1, \ldots, m$ while by the outer induction hypothesis $D(t_{j,\nu_{j,k-1}}^{(k-1)}, 1, 1) \leq k-1$ for $j = 1, \ldots, m$. Thus, each term in (3.3.25) is of the form required for (3.3.18), completing the proof that (3.3.18) holds for k and $\rho(t) = \bar{\rho}$. Hence, by the inner induction, (3.3.18) holds for k and all $t \in T$. So, by the outer induction, (3.3.18) holds for all $k \geq 1$ and all $t \in T$, completing the proof of part (a) of the lemma.

To prove part (b) of the lemma, note that (3.3.1) holds for $k \geq 0$, since (3.1.14) is satisfied with $q \geq 0$. Moreover, we showed above that (3.3.18) holds for $k \geq 1$. Therefore, for $k \geq 2$ and $t = [t_1, \ldots, t_m] \in T$ with $\rho(t) \geq 1$, we can substitute the right side of (3.3.18) for $\phi^{(k)}(t)$, $\phi^{(k)}(t_i)$ and $\phi^{(k-1)}(t_i)$ into (3.3.1) to get

$$\begin{aligned} \sum_{\nu=1}^{\mathcal{N}_{t,k}} \kappa_\nu^{(k)}(t) \phi(t_\nu^{(k)}) &= \rho(t) A \prod_{i=1}^m \left[\sum_{\nu=1}^{\mathcal{N}_{t_i,k-1}} \kappa_\nu^{(k-1)}(t_i) \phi(t_{i,\nu}^{(k-1)})\right] \\ &+ \rho(t) A \sum_{i=1}^m \left(\prod_{\substack{j=1 \\ j \neq i}}^m \left[\sum_{\nu=1}^{\mathcal{N}_{t_j,k-1}} \kappa_\nu^{(k-1)}(t_j) \phi(t_{j,\nu}^{(k-1)})\right]\right) \\ &\times \left(\left[\sum_{\nu=1}^{\mathcal{N}_{t_i,k}} \kappa_\nu^{(k)}(t_i) \phi(t_{i,\nu}^{(k)})\right] - \left[\sum_{\nu=1}^{\mathcal{N}_{t_i,k-1}} \kappa_\nu^{(k-1)}(t_i) \phi(t_{i,\nu}^{(k-1)})\right]\right) \end{aligned} \quad (3.3.26)$$

Matching terms in (3.3.26). we get

$$\mathcal{K}^{(k)}(t) = \prod_{i=1}^{m} \mathcal{K}^{(k-1)}(t_i) + \sum_{i=1}^{m} \left(\prod_{\substack{j=1\\ j\neq i}}^{m} \mathcal{K}^{(k-1)}(t_j) \right) \left(\mathcal{K}^{(k)}(t_i) - \mathcal{K}^{(k-1)}(t_i) \right). \tag{3.3.27}$$

Moreover, if we let $\mathcal{K}^{(0)}(t) = \zeta(t)$ for all $t \in T$, then, using a similar argument, we see that (3.3.27) also holds for $k = 1$ and all $t = [t_1, \ldots, t_m] \in T$ with $\rho(t) \geq 1$.

To prove (3.3.19), we use (3.3.27) together with an inner and an outer induction: the outer induction is on k while the inner induction, used to prove both the base case ($k = 1$) and the induction step ($k \geq 2$) of the outer induction, is on $\rho(t)$.

To prove (3.3.19) for $k = 1$, assume first that $\rho(t) = 0$. Therefore, $t = \emptyset$ and $\phi^{(1)}(\emptyset) = \varepsilon_s = \phi(\emptyset)$ by Lemma 3.3.1 and Theorem 2.4. Therefore, $\mathcal{N}_{\emptyset,1} = 1$ and $\kappa_1^{(1)}(\emptyset) = 1$. So, $\mathcal{K}^{(1)}(\emptyset) = \kappa_1^{(1)}(\emptyset) = 1$ as required.

Next assume (3.3.19) holds for $k = 1$ but with the additional restriction that $\rho(t) < \bar{\rho}$ for some $\bar{\rho} \geq 1$. Let $t = [t_1, \ldots, t_m]$ be any tree for which $D(t, q, q) \leq 1$ and $\rho(t) = \bar{\rho}$. Therefore, $\rho(t_i) < \bar{\rho}$ for $i = 1, \ldots, m$ and, after possibly re-ordering the subtrees t_i, $D(t_i, q, q) = 0$ for $i = 1, \ldots, m-1$ and $D(t_m, q, q) \leq 1$ by Lemma 2.8. Since $D(t_i, q, q) = 0$ implies $\rho(t_i) \leq q$ by Definition 2.5, it follows from (3.1.14) that $\mathcal{K}^{(0)}(t_i) = \zeta(t_i) = 1$ for $i = 1, \ldots, m-1$. Moreover, by the inner induction hypothesis on $\rho(t)$, $\mathcal{K}^{(1)}(t_i) = 1$ for $i = 1, \ldots, m$. Hence, by (3.3.27),

$$\begin{aligned}
\mathcal{K}^{(1)}(t) &= \prod_{i=1}^{m} \mathcal{K}^{(0)}(t_i) + \sum_{i=1}^{m} \left(\prod_{\substack{j=1\\ j\neq i}}^{m} \mathcal{K}^{(0)}(t_j) \right) \left(\mathcal{K}^{(1)}(t_i) - \mathcal{K}^{(0)}(t_i) \right) \\
&= \prod_{i=1}^{m} \mathcal{K}^{(0)}(t_i) + \left(\prod_{j=1}^{m-1} \mathcal{K}^{(0)}(t_j) \right) \left(\mathcal{K}^{(1)}(t_m) - \mathcal{K}^{(0)}(t_m) \right) \\
&= \left(\prod_{i=1}^{m-1} \mathcal{K}^{(0)}(t_i) \right) \mathcal{K}^{(1)}(t_m) \\
&= 1.
\end{aligned}$$

Therefore, (3.3.19) holds for $k = 1$, completing the proof of the base case of the outer induction on k.

For the outer induction step on k, let $k \geq 2$ and assume that (3.3.19) holds for $k-1$. We use an inner induction on $\rho(t)$ to show (3.3.19) holds for k.

If $\rho(t) = 0$, then $t = \emptyset$ and $\phi^{(k)}(\emptyset) = \varepsilon_s = \phi(\emptyset)$ by Lemma 3.3.1 and Theorem 2.4. Therefore, $\mathcal{N}_{\emptyset,k} = 1$ and $\kappa_1^{(k)}(\emptyset) = 1$. So, $\mathcal{K}^{(k)}(\emptyset) = \kappa_1^{(k)}(\emptyset) = 1$, as required.

For the inner induction step on $\rho(t)$, assume also that (3.3.19) holds for k but with the additional restriction that $\rho(t) < \bar{\rho}$ for some $\bar{\rho} \geq 1$. Let $t = [t_1, \ldots, t_m]$ be any tree satisfying $D(t, q, q) \leq k$ and $\rho(t) = \bar{\rho}$. Therefore, $\rho(t_i) < \bar{\rho}$ for $i = 1, \ldots, m$ and, after possibly re-ordering the subtrees t_i, $D(t_i, q, q) \leq k-1$ for $i = 1, \ldots, m-1$ and $D(t_m, q, q) \leq k$ by Lemma 2.8. So, by the outer induction hypothesis on k, $\mathcal{K}^{(k-1)}(t_i) = 1$ for $i = 1, \ldots, m-1$.

Moreover, by the inner induction hypothesis on $\rho(t)$, $\mathcal{K}^{(k)}(t_i) = 1$ for $i = 1, \ldots, m$. So by (3.3.27)

$$\begin{aligned}
\mathcal{K}^{(k)}(t) &= \prod_{i=1}^{m} \mathcal{K}^{(k-1)}(t_i) + \sum_{i=1}^{m} \left(\prod_{\substack{j=1 \\ j \neq i}}^{m} \mathcal{K}^{(k-1)}(t_j) \right) \left(\mathcal{K}^{(k)}(t_i) - \mathcal{K}^{(k-1)}(t_i) \right) \\
&= \prod_{i=1}^{m} \mathcal{K}^{(k-1)}(t_i) + \left(\prod_{j=1}^{m-1} \mathcal{K}^{(k-1)}(t_j) \right) \left(\mathcal{K}^{(k)}(t_m) - \mathcal{K}^{(k-1)}(t_m) \right) \\
&= \left(\prod_{i=1}^{m-1} \mathcal{K}^{(k-1)}(t_i) \right) \mathcal{K}^{(k)}(t_m) \\
&= 1.
\end{aligned}$$

Therefore, (3.3.19) holds for k, completing the proof of part (b) of the lemma.

We prove part (c) of the lemma by induction on k. Let $\hat{t}$ be any tree such $\rho(\hat{t}) = q + 1$ and $\zeta(\hat{t}) \neq 1$. Such a $\hat{t}$ exists, since q is the largest integer for which (3.1.14) holds.

To see that (3.3.20) holds for $k = 1$, recall that (3.3.27) holds for $k = 1$ with $\mathcal{K}^{(0)}(t) = \zeta(t)$. In addition, since $\rho(\hat{t}) = q + 1$, $D(\hat{t}, q, q) = 1$, whence $\mathcal{K}^{(1)}(\hat{t}) = 1$ by (3.3.19). So

$$\begin{aligned}
\mathcal{K}^{(1)}(t_{D,1}) &= \mathcal{K}^{(1)}([\hat{t}, \hat{t}]) \\
&= \left[\mathcal{K}^{(0)}(\hat{t})\right]^2 + 2\mathcal{K}^{(0)}(\hat{t}) \left[\mathcal{K}^{(1)}(\hat{t}) - \mathcal{K}^{(0)}(\hat{t})\right] \\
&= \left[\zeta(\hat{t})\right]^2 + 2\zeta(\hat{t}) \left[1 - \zeta(\hat{t})\right] \\
&= 1 - \left[1 - \zeta(\hat{t})\right]^2 \\
&\neq 1.
\end{aligned}$$

For the induction step, let $k \geq 2$ and assume that (3.3.20) holds for $k - 1$. That is, $\mathcal{K}^{(k-1)}(t_{D,k-1}) = 1 - \left(1 - \zeta(\hat{t})\right)^{2^{k-1}}$. Note also that it is easy to prove by induction on $k \geq 2$ that $D(t_{D,k-1}, q, q) = k$, from which it follows that $\mathcal{K}^{(k)}(t_{D,k-1}) = 1$ by (3.3.19). So by (3.3.27)

$$\begin{aligned}
\mathcal{K}^{(k)}(t_{D,k}) &= \mathcal{K}^{(k)}([t_{D,k-1}, t_{D,k-1}]) \\
&= \left[\mathcal{K}^{(k-1)}(t_{D,k-1})\right]^2 + 2\mathcal{K}^{(k-1)}(t_{D,k-1}) \left(\mathcal{K}^{(k)}(t_{D,k-1}) - \mathcal{K}^{(k-1)}(t_{D,k-1}) \right) \\
&= \left[1 - \left(1 - \zeta(\hat{t})\right)^{2^{k-1}}\right]^2 + 2\left[1 - \left(1 - \zeta(\hat{t})\right)^{2^{k-1}}\right] \left(1 - \left[1 - \left(1 - \zeta(\hat{t})\right)^{2^{k-1}}\right]\right) \\
&= 1 - 2\left[1 - \zeta(\hat{t})\right]^{2^{k-1}} + \left[1 - \zeta(\hat{t})\right]^{2^k} + 2\left[1 - \zeta(\hat{t})\right]^{2^{k-1}} - 2\left[1 - \zeta(\hat{t})\right]^{2^k} \\
&= 1 - \left[1 - \zeta(\hat{t})\right]^{2^k} \\
&\neq 1,
\end{aligned}$$

completing the proof of part (c) of the lemma. □

Proof of Theorem 3.3.4. We first show that, for $k \geq 1$,

$$\tilde{\phi}^{(k)}_{s+1}(t) = 1 \quad \forall\, t \in T \text{ s.t. } \rho(t) \leq \min(p^D_k, \bar{q}_k). \tag{3.3.28}$$

By Theorem 2.1, this ensures that the order of $y^{(k)}_{n+1}$ is at least $\min(p^D_k, \bar{q}_k)$.

To begin, note that $\tilde{\phi}^{(k)}_{s+1}(\emptyset) = 1$ for $k \geq 1$ by Lemma 3.3.1, so (3.3.28) holds for $t = \emptyset$. Therefore, assume $t \neq \emptyset$ for the remainder of the proof of (3.3.28).

To prove (3.3.28) holds for $k = 1$ and $t \neq \emptyset$, let $t = [t_1, \ldots, t_m]$ be any tree for which $1 \leq \rho(t) \leq \min(p^D_1, \bar{q}_1)$. Since $\rho(t) \leq \bar{q}_1 = q + 1$, $\rho(t_i) \leq q$ for $i = 1, \ldots, m$, whence $\phi^{(0)}(t_i) = c^{\rho(t_i)}$ for $i = 1, \ldots, m$ by (3.1.14). In addition, note $\sum_{i=1}^m \rho(t_i) = \rho(t) - 1$. Moreover, $\phi^{(0)}(\emptyset) = \varepsilon_s$, since $q \geq 0$. Therefore, we can apply (3.3.2) together with Theorem 2.4 and the relations noted above to yield

$$\tilde{\phi}^{(1)}_{s+1}(t) = \rho(t) b^T \prod_{i=1}^m \phi^{(0)}(t_i) = \rho(t) b^T \prod_{i=1}^m c^{\rho(t_i)} = \rho(t) b^T c^{\rho(t)-1} = \tilde{\phi}_{s+1}([\tau^{\rho(t)-1}]). \tag{3.3.29}$$

In addition, since $\rho([\tau^{\rho(t)-1}]) = \rho(t) \leq p^D_1$ and $D(\tau, 1, 1) = 0$,

$$\tilde{\phi}^{(1)}_{s+1}(t) = \tilde{\phi}_{s+1}([\tau^{\rho(t)-1}]) = 1$$

from (3.3.29) and (3.3.14). Therefore, (3.3.28) holds for $k = 1$.

To prove (3.3.28) for $k \geq 2$ and $t \neq \emptyset$, let $t = [t_1, \ldots, t_m]$ be any tree satisfying $1 \leq \rho(t) \leq \min(p^D_k, \bar{q}_k)$. Since $\rho(t) \leq \bar{q}_k = 2^{k-1}(q + 2) - 1$, $\rho(t_i) \leq 2^{k-1}(q + 2) - 2$ for $i = 1, \ldots, m$, whence $D(t_i, q, q) \leq k - 1$ for $i = 1, \ldots, m$ by Lemma 2.9. From part (a) of Lemma 3.3.5, we get that

$$\phi^{(k-1)}(t_i) = \sum_{\nu=1}^{\mathcal{N}_{t_i,k-1}} \kappa^{(k-1)}_\nu(t_i) \phi(t^{(k-1)}_{i,\nu}) \tag{3.3.30}$$

where each $t^{(k-1)}_{i,\nu}$ is a tree satisfying $\rho(t^{(k-1)}_{i,\nu}) = \rho(t_i)$ and $D(t^{(k-1)}_{i,\nu}, 1, 1) \leq k - 1$. Moreover, since $D(t_i, q, q) \leq k - 1$ for $i = 1, \ldots, m$, it follows from part (b) of Lemma 3.3.5 that

$$\mathcal{K}^{(k-1)}(t_i) \equiv \sum_{\nu=1}^{\mathcal{N}_{t_i,k-1}} \kappa^{(k-1)}_\nu(t_i) = 1 \tag{3.3.31}$$

for $i = 1, \ldots, m$. In addition, since $k \geq 2$, $\tilde{\phi}^{(k-1)}(\emptyset) = \varepsilon_{s+1}$ by Lemma 3.3.1. So, we can apply (3.3.2) and (3.3.30) to yield

$$\begin{aligned}
\tilde{\phi}^{(k)}_{s+1}(t) &= \rho(t) b^T \prod_{i=1}^m \phi^{(k-1)}(t_i) \\
&= \rho(t) b^T \prod_{i=1}^m \left(\sum_{\nu=1}^{\mathcal{N}_{t_i,k-1}} \kappa^{(k-1)}_\nu(t_i) \phi(t^{(k-1)}_{i,\nu}) \right) \qquad (3.3.32) \\
&= \sum_{\nu=1}^{\tilde{\mathcal{N}}_{t,k}} \tilde{\kappa}^{(k)}_\nu(t) \tilde{\phi}_{s+1}(t^{(k)}_\nu)
\end{aligned}$$

where

$$\sum_{\nu=1}^{\tilde{\mathcal{N}}_{t,k}} \tilde{\kappa}_\nu^{(k)}(t) = \prod_{i=1}^{m} \left(\sum_{\nu=1}^{\mathcal{N}_{t_i,k-1}} \kappa_\nu^{(k-1)}(t_i) \right) = 1 \tag{3.3.33}$$

by (3.3.31). In addition, note that each $t_\nu^{(k)} = [t_{1,\nu_1}^{(k-1)}, \dots, t_{m,\nu_m}^{(k-1)}]$ with each $t_{i,\nu_i}^{(k-1)}$ one of $t_{i,\nu}^{(k-1)}$ for $\nu = 1, \dots, \mathcal{N}_{t_i,k-1}$ and $i = 1, \dots, m$. Since $t = [t_1, \dots, t_m]$ and each $\rho(t_{i,\nu_i}^{(k-1)}) = \rho(t_i)$, $\rho(t_\nu^{(k)}) = \rho(t) \leq p_k^D$ for $\nu = 1, \dots, \tilde{\mathcal{N}}_{t,k}$. Moreover, each $D(t_{i,\nu}^{(k-1)}, 1, 1) \leq k-1$. Therefore, each $\tilde{\phi}_{s+1}(t_\nu^{(k)}) = 1$ by (3.3.14). Combining this with (3.3.32) and (3.3.33), we get

$$\tilde{\phi}_{s+1}^{(k)}(t) = \sum_{\nu=1}^{\tilde{\mathcal{N}}_{t,k}} \tilde{\kappa}_\nu^{(k)}(t) = 1$$

completing the proof of (3.3.28).

To complete the proof, we need to show that, if either $q \geq 1$ or $\bar{q}_k < p_k^D$, then the order of $y_{n+1}^{(k)}$ is at most $\min(p_k^D, \bar{q}_k)$, since this together with (3.3.28) ensures that the order of $y_{n+1}^{(k)}$ is exactly $\min(p_k^D, \bar{q}_k)$.

Assume first that $\bar{q}_k < p_k^D$. We split the proof for this case into two subcases: $k = 1$ and $k \geq 2$. For $k = 1$, let $\hat{t}$ be any tree for which $\rho(\hat{t}) = q+1$ and $\zeta(\hat{t}) \neq 1$. Such a $\hat{t}$ exists, since q is the largest integer for which (3.1.14) holds. Since (3.1.14) holds with $q \geq 0$, $\phi^{(0)}(\emptyset) = \varepsilon_s$, so we may apply (3.3.2), which together with (3.1.14) and Theorem 2.4 gives

$$\tilde{\phi}_{s+1}^{(1)}([\hat{t}]) = \rho([\hat{t}]) b^T \phi^{(0)}(\hat{t}) = \rho([\hat{t}]) b^T \zeta(\hat{t}) c^{\rho(\hat{t})} = \zeta(\hat{t}) \tilde{\phi}_{s+1}([\tau^{\rho(\hat{t})}]). \tag{3.3.34}$$

Since $\rho(\hat{t}) = q+1 = \bar{q}_1 < p_1^D$, $\rho([\tau^{\rho(\hat{t})}]) = \rho([\hat{t}]) = q+2 = \bar{q}_1 + 1 \leq p_1^D$. Moreover, since $D(\tau, 1, 1) = 0$, $\tilde{\phi}_{s+1}([\tau^{\rho(\hat{t})}]) = 1$ by (3.3.14). Combining this with (3.3.34), we get $\tilde{\phi}_{s+1}^{(1)}([\hat{t}]) = \zeta(\hat{t}) \neq 1$. Since, as noted above, $\rho([\hat{t}]) = \bar{q}_1 + 1$, it follows from Theorem 2.1 that the order of $y_{n+1}^{(1)}$ is at most $\bar{q}_1 = \min(p_1^D, \bar{q}_1)$, as required.

Next assume $k \geq 2$ and $\bar{q}_k < p_k^D$. Let $t_{D,k}$ be defined as in part (c) of Lemma 3.3.5. By part (a) of Lemma 3.3.5,

$$\phi^{(k-1)}(t_{D,k-1}) = \sum_{\nu=1}^{\mathcal{N}_{t_{D,k-1},k-1}} \kappa_\nu^{(k-1)}(t_{D,k-1}) \phi(t_{D,k-1,\nu}^{(k-1)}) \tag{3.3.35}$$

where each $t_{D,k-1,\nu}^{(k-1)}$ is a tree satisfying $\rho(t_{D,k-1,\nu}^{(k-1)}) = \rho(t_{D,k-1})$ and $D(t_{D,k-1,\nu}^{(k-1)}, 1, 1) \leq k-1$. Moreover, it is easy to show by induction on k that $\rho(t_{D,k-1}) = 2^{k-1}(q+2) - 1 = \bar{q}_k$. So $\rho([t_{D,k-1,\nu}^{(k-1)}]) = \rho([t_{D,k-1}]) = \bar{q}_k + 1 \leq p_k^D$. Combine this with Theorem 2.4 and (3.3.14) to get

$$\rho([t_{D,k-1}]) b^T \phi(t_{D,k-1,\nu}^{(k-1)}) = \rho([t_{D,k-1,\nu}^{(k-1)}]) b^T \phi(t_{D,k-1,\nu}^{(k-1)}) = \tilde{\phi}_{s+1}([t_{D,k-1,\nu}^{(k-1)}]) = 1. \tag{3.3.36}$$

Moreover, by part (c) of Lemma 3.3.5,

$$\mathcal{K}^{(k-1)}(t_{D,k-1}) \equiv \sum_{\nu=1}^{\mathcal{N}_{t_{D,k-1},k-1}} \kappa_\nu^{(k-1)}(t_{D,k-1}) \neq 1. \tag{3.3.37}$$

Since $\phi^{(k-1)}(\emptyset) = \varepsilon_s$ by Lemma 3.3.1, we can apply (3.3.2) together with (3.3.35), (3.3.36) and (3.3.37) to get

$$\begin{aligned}
\tilde{\phi}^{(k)}_{s+1}([t_{D,k-1}]) &= \rho([t_{D,k-1})b^T\phi^{(k-1)}(t_{D,k-1}) \\
&= \rho([t_{D,k-1}])b^T \sum_{\nu=1}^{\mathcal{N}_{t_{D,k-1},k-1}} \kappa_\nu^{(k-1)}(t_{D,k-1})\phi(t^{(k-1)}_{D,k-1,\nu}) \\
&= \sum_{\nu=1}^{\mathcal{N}_{t_{D,k-1},k-1}} \kappa_\nu^{(k-1)}(t_{D,k-1})\left(\rho([t_{D,k-1}])b^T\phi(t^{(k-1)}_{D,k-1,\nu})\right) \\
&= \sum_{\nu=1}^{\mathcal{N}_{t_{D,k-1},k-1}} \kappa_\nu^{(k-1)}(t_{D,k-1}) \\
&\neq 1.
\end{aligned}$$

Since, as noted above, $\rho([t_{D,k-1}]) = \bar{q}_k + 1$, it follows from Theorem 2.1 that the order of $y^{(k)}_{n+1}$ is at most $\bar{q}_k = \min(p^D_k, \bar{q}_k)$, as required.

It remains to show that the order of $y^{(k)}_{n+1}$ is at most $\min(p^D_k, \bar{q}_k)$ if $q \geq 1$. However, we have just shown that this is true if $q \geq 0$ and $p^D_k > q_k$. Therefore, to complete the proof, we need only consider the case $q \geq 1$ and $p^D_k \leq q_k$.

From (3.3.14), the definition of p^D_k, there exists a tree $\hat{t} = [\hat{t}_1, \ldots, \hat{t}_m]$ for which $\rho(\hat{t}) = p^D_k + 1$, $D(\hat{t}_i, 1, 1) \leq k-1$ for $i = 1, \ldots, m$ and $\tilde{\phi}_{s+1}(\hat{t}) \neq 1$.

If $k = 1$, then $D(\hat{t}_i, 1, 1) \leq k-1$ implies $D(\hat{t}_i, 1, 1) = 0$, whence $\rho(\hat{t}_i) \leq 1$ for $i = 1, \ldots, m$ by Definition 2.5. Therefore, $\phi^{(0)}(\hat{t}_i) = c^{\rho(\hat{t}_i)}$ for $i = 1, \ldots, m$ by (3.1.14) since $q \geq 1$. In addition, note that the only trees for which $\rho(t) \leq 1$ are $\emptyset$ and τ. Moreover, $\phi(\emptyset) = c^0$ and $\phi(\tau) = c^1$ by Theorem 2.4. Therefore, $\phi^{(0)}(\hat{t}_i) = \phi(\hat{t}_i)$ for $i = 1, \ldots, m$.

If $k \geq 2$, note that, since $D(\hat{t}_i, 1, 1) \leq k-1$ for $i = 1, \ldots, m$ and $q \geq 1$, we may take $r = 1$ in (3.3.6) to conclude that $\phi^{(k-1)}(\hat{t}_i) = \phi(\hat{t}_i)$ for $i = 1, \ldots, m$.

Thus we have established that, for $k \geq 1$, $\phi^{(k-1)}(\hat{t}_i) = \phi(\hat{t}_i)$ for $i = 1, \ldots, m$. Moreover, $\phi^{(0)}(\emptyset) = \varepsilon_s$ since $q \geq 1$ and $\phi^{(k)}(\emptyset) = \varepsilon_s$ for $k \geq 1$ by Lemma 3.3.1. Consequently, for $k \geq 1$, we can apply (3.3.2) and Theorem 2.4 to yield

$$\tilde{\phi}^{(k)}_{s+1}(\hat{t}) = \rho(\hat{t})b^T \prod_{i=1}^{m} \phi^{(k-1)}(\hat{t}_i) = \rho(\hat{t})b^T \prod_{i=1}^{m} \phi(\hat{t}_i) = \tilde{\phi}_{s+1}(\hat{t}) \neq 1.$$

Since, as noted above, $\rho(\hat{t}) = p^D_k + 1$, it follows from Theorem 2.1 that the order of $y^{(k)}_{n+1}$ is at most $p^D_k = \min(p^D_k, \bar{q}_k)$, as required. $\square$

4 Numerical Results.

In this section, we present some numerical experiments which confirm the theoretical results given in §3. In each experiment, we solve the problem

$$\begin{aligned} y_1' &= 4xy_1 + 2\sqrt{y_1}\, y_2, & y_1(x_0) &= e^{2x_0^2}\sin^2(x_0), \\ y_2' &= 2xy_2 + y_1 + y_2^2 - e^{x^2}\left(\sin(x) + e^{x^2}\right), & y_2(x_0) &= e^{x_0^2}\cos(x_0), \end{aligned} \tag{4.1}$$

using MATLAB on a Sun Sparc-station IPX computer. In each case, we integrated from $x_0 = 0.2$ to $x_{\text{end}} = 1.2$ with the iterated Runge-Kutta method (1.13)–(1.14). We experimented with several underlying RK formulas together with different predictors to compute the starting values $\{Y_{n,i}^{(0)}\}$ for the iteration. The schemes used in the experiments are chosen partly to illustrate rather regular order behaviour (Examples 4.1 and 4.2) as well as more surprising results (Examples 4.3, 4.4 and 4.5) that are difficult to explain without the use of B-series.

We used a fixed stepsize h and the same number of iterations k on each step throughout any one numerical integration of (4.1), but we varied h and k from integration to integration. Since the true solution of this problem is

$$y_1(x) = e^{2x^2}\sin^2(x), \qquad y_2(x) = e^{x^2}\cos(x),$$

we are able to compute the global error for several integrations with the same k but different h's to calculate an *observed* order for the iterated RK method. The global errors and the observed orders of component y_1 are listed in Tables 4.1–4.5. The predicted order in each case is $\min(p, q_k)$, where p is the order of the underlying method and q_k is given by Theorem 1.1.

In the tables below, we use Fortran notation for some numbers. For example, we represent $6.25 \cdot 10^{-3}$ as 6.25e–3.

Example 4.1 The iterated RK method in this example is based on the 4-stage Kuntzmann-Butcher formula [8, p. 205] of order $p = 8$ together with the trivial predictor $Y_{n,i}^{(0)} = y_n$ for $i = 1, \ldots, s$, whence $q = r = 0$. The methods use simple, modified Newton and full Newton iteration. The numerical results given in Table 4.1 show that the predicted and observed orders agree well.

Example 4.2 As in Example 4.1, the underlying method is the 4-stage Kuntzmann-Butcher formula [8, p. 205] of order $p = 8$, but, in this case, we use the associated collocation polynomial to calculate $Y_n^{(0)}$. Thus, for $k \geq 2$, $q = r = 4$. For $k = 1$ no iterations are done on the stage values, thus $Y_n^{(0)}$ is based on the stage values from the first step, resulting in an ineffective method. The numerical results given in Table 4.2 again show that the theoretical and observed orders agree well.

Example 4.3 The underlying RK formula is the 8^{th}-order extrapolation method based on the implicit midpoint rule written as the 10-stage diagonally-implicit RK formula shown in Figure 4.1. The starting values $Y_n^{(0)}$ for the iteration are computed by Heun's 3$^{\text{rd}}$-order explicit RK formula [8, p. 133], whence $q = 3$. However, the stage order of the underlying RK formula is 1, so $r = 1$.

We chose this scheme to show that an iterated RK method can benefit from a high-order predictor, even if the stage order of the underlying RK formula is low. The numerical results given in Table 4.3 show that the observed orders for the simple and modified Newton iterations agree well with the orders predicted by Theorem 1.1, but, for full Newton iteration with $k = 2$, the observed order is 8, while $q_2 = 7$. However, the predictor satisfies the stronger assumption (3.1.14), so Theorem 3.3.4 can be used to ensure that the order of $y_{n+1}^{(2)}$ is exactly $\min(p_2^D, \bar{q}_2) = \min(8, 9) = 8$. Thus, the observed and the theoretical orders do agree.

We did not list the numerical results for $k = 4$ for the full Newton iteration in Table 4.3, as including them makes the table too wide. However, the observed order is approximately 8 in this case, in good agreement with the theoretical results.

$$
\begin{array}{c|cccccccccc}
\frac{1}{2} & \frac{1}{2} \\
\frac{1}{4} & 0 & \frac{1}{4} \\
\frac{3}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\
\frac{1}{6} & 0 & 0 & 0 & \frac{1}{6} \\
\frac{1}{2} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{6} \\
\frac{5}{6} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \\
\frac{1}{8} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{8} \\
\frac{3}{8} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{8} \\
\frac{5}{8} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} \\
\frac{7}{8} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} \\
\hline
 & -\frac{1}{360} & \frac{8}{45} & \frac{8}{45} & -\frac{243}{280} & -\frac{243}{280} & -\frac{243}{280} & \frac{256}{315} & \frac{256}{315} & \frac{256}{315} & \frac{256}{315}
\end{array}
$$

Figure 4.1: The 8$^{\text{th}}$-order extrapolation method based on the implicit midpoint rule written as a 10-stage diagonally-implicit RK formula.

Example 4.4 In this example, we illustrate the advantage enjoyed by iterated RK methods based on predictors that satisfy (3.1.14). We again use the 8$^{\text{th}}$-order extrapolation scheme exhibited in Figure 4.1 as the underlying RK formula, but we replace the 3$^{\text{rd}}$-order predictor used in the previous example by Heun's 2$^{\text{nd}}$-order explicit RK formula [4, p. 173]. We refer to this iterated RK method as Scheme A. Note, $q = 2$ and $r = 1$. Only one full Newton iteration is performed, so $k = 2$, $q_2 = 6$ and $\bar{q}_2 = 7$. As in the preceding example, the assumptions of Theorem 3.3.4 are satisfied and the observed order in Table 4.4 for this case agrees well with the predicted order $\bar{q}_2 = 7$, and is clearly greater than the order $q_2 = 6$ ensured by Theorem 1.1.

Next we modify the iterated RK method described above. We again use the 8^{th}-order extrapolation scheme exhibited in Figure 4.1 as the underlying RK formula, but now we use Heun's 2^{nd}-order explicit RK formula [4, p. 173] to predict the internal stage values $Y_{n,i}^{(0)}$ for $i = 1, 2, 4, 5, 7$ and 8 only. (The c_i associated with each of these $Y_{n,i}^{(0)}$s satisfies $c_i \in [0, 1/2]$.) We use Heun's more accurate 3^{rd}-order explicit RK formula [8, p. 133] employed in Example 4.3 to predict the other internal stage values $Y_{n,i}^{(0)}$ for $i = 3, 6, 9$ and 10. (The c_i associated with each of these $Y_{n,i}^{(0)}$s satisfies $c_i \in (1/2, 1]$.) We refer to this iterated RK method as Scheme B. Note that $q = 2$ and $r = 1$ for it too. Moreover, as with Scheme A, only one full Newton iteration is performed in Scheme B, so $k = 2$ and $q_2 = 6$. However, unlike Scheme A, the predictor associated with Scheme B does not satisfy (3.1.14).

Since Scheme B's predictor is of higher order than Scheme A's for the $Y_{n,i}^{(0)}$s in the second half of the step, one might expect the former method to be more accurate than the latter. The numerical results in Table 4.4 show that this is indeed the case for the larger values of h listed there. However, for the smaller values of h, Scheme A is more accurate than Scheme B. Moreover, as noted above, the observed order for Scheme A is $\bar{q}_2 = 7$, while the observed order for Scheme B is about $q_2 = 6$, in good agreement with the order predicted by Theorem 1.1, but distinctly less than $\bar{q}_2 = 7$ attained by Scheme A which satisfies the assumptions of Theorem 3.3.4.

Example 4.5 This example demonstrates the possibly surprising result that the order of $y_{n+1}^{(k)}$ may decrease as k increases. The underlying RK method is the 4-stage 3^{rd}-order DIRK formula shown in Figure 4.2. The H-restricted order values for this formula are $p_1^H = 6$, $p_2^H = 4$ and $p_k^H = 3$ for $k \geq 3$. The starting values $Y_n^{(0)}$ for the iteration are computed by Fehlberg's 8^{th}-order explicit RK-formula, [8, p. 194]. So $q = 8$. Moreover, the iterated RK method employs simple iteration. By Theorem 3.1.4 the order of $y_{n+1}^{(k)}$ is p_k^H. For completeness, the numerical results for the modified and full Newton iterations are included also in Table 4.5. Note that these results are in complete agreement with our theory.

$$
\begin{array}{c|cccc}
\frac{1}{2}+\frac{\sqrt{5}}{10} & \frac{1}{2}+\frac{\sqrt{5}}{10} & & & \\
0 & \frac{5}{4}+\frac{3\sqrt{5}}{4} & -\frac{5}{4}-\frac{3\sqrt{5}}{4} & & \\
\frac{1}{2}-\frac{\sqrt{5}}{10} & -\frac{1}{3}-\frac{11\sqrt{5}}{30} & \frac{5}{12}+\frac{3\sqrt{5}}{20} & \frac{5}{12}+\frac{7\sqrt{5}}{60} & \\
1 & \frac{5}{12} & \frac{1}{12} & \frac{5}{12} & \frac{1}{12} \\
\hline
1 & \frac{5}{12} & \frac{1}{12} & \frac{5}{12} & \frac{1}{12}
\end{array}
$$

Figure 4.2: The 4-stage 3^{rd}-order DIRK formula used in Example 4.5.

Simple Iteration.

	$k = 1,\ q_1 = 1$		$k = 2,\ q_2 = 2$		$k = 3,\ q_3 = 3$		$k = 4,\ q_4 = 4$	
h	Error	Order	Error	Order	Error	Order	Error	Order
1.00e–1	1.02e+1		3.14e+0		5.12e–1		6.08e–2	
5.00e–2	7.61e+0	0.42	1.09e+0	1.53	8.12e–2	2.66	4.75e–3	3.68
2.50e–2	5.07e+0	0.58	3.15e–1	1.79	1.13e–2	2.84	3.31e–4	3.84
1.25e–2	3.06e+0	0.73	8.38e–2	1.91	1.49e–3	2.92	2.19e–5	3.92
6.25e–3	1.71e+0	0.84	2.14e–2	1.96	1.91e–4	2.96	1.41e–6	3.96

Modified Newton Iteration.

	$k = 1,\ q_1 = 1$		$k = 2,\ q_2 = 3$		$k = 3,\ q_3 = 5$		$k = 4,\ q_4 = 7$	
h	Error	Order	Error	Order	Error	Order	Error	Order
1.00e–1	1.02e+1		5.84e–2		9.33e–4		8.71e–6	
5.00e–2	7.61e+0	0.42	9.54e–3	2.61	2.88e–5	5.02	6.79e–8	7.00
2.50e–2	5.07e+0	0.58	1.35e–3	2.82	8.77e–7	5.04	4.94e–10	7.10
1.25e–2	3.06e+0	0.73	1.80e–4	2.91	2.69e–8	5.03	3.52e–12	7.13
6.25e–3	1.71e+0	0.84	2.31e–5	2.96	8.32e–10	5.02	3.55e–14	6.63

Full Newton Iteration.

	$k = 1,\ q_1 = 1$		$k = 2,\ q_2 = 3$		$k = 3,\ q_3 = 7$		$k = 4,\ q_4 = 15$	
h	Error	Order	Error	Order	Error	Order	Error	Order
1.00e–1	1.02e+1		9.01e–2		7.99e–6		6.10e–8	
5.00e–2	7.61e+0	0.42	1.15e–2	2.96	7.43e–8	6.75	2.43e–10	7.97
2.50e–2	5.07e+0	0.58	1.46e–3	2.98	6.17e–10	6.91	1.08e–12	7.81
1.25e–2	3.06e+0	0.73	1.84e–4	2.99	4.77e–12	7.02	3.55e–15	8.25
6.25e–3	1.71e+0	0.84	2.30e–5	3.00	3.73e–14	7.00	3.37e–14	–3.25

Table 4.1: Results for Example 4.1 based on the 4-stage Kuntzmann-Butcher formula of order $p = 8$ using the trivial predictor $Y_{n,i}^{(0)} = y_n$ for $i = 1, \ldots, s$, whence $q = r = 0$.

Simple Iteration.

	$k=2,\ q_2=6$		$k=3,\ q_3=7$		$k=4,\ q_4=8$	
h	Error	Order	Error	Order	Error	Order
1.00e–1	5.09e–2		5.66e–4		2.84e–5	
5.00e–2	5.81e–4	6.45	4.30e–6	7.04	1.55e–7	7.52
2.50e–2	4.75e–6	6.93	3.49e–8	6.94	7.42e–10	7.71
1.25e–2	4.66e–8	6.67	2.89e–10	6.92	3.20e–12	7.86
6.25e–3	5.48e–10	6.41	2.42e–12	6.90	1.46e–13	7.43

Modified Newton Iteration.

	$k=2,\ q_2=7$		$k=3,\ q_3=9$		$k=4,\ q_4=11$	
h	Error	Order	Error	Order	Error	Order
1.00e–1	3.54e–5		2.70e–7		6.07e–8	
5.00e–2	2.54e–7	7.12	1.24e–9	7.77	2.41e–10	7.97
2.50e–2	1.33e–9	7.58	4.29e–12	8.17	1.04e–12	7.86
1.25e–2	5.81e–12	7.83	3.73e–14	6.85	3.55e–15	8.20
6.25e–3	3.38e–14	7.43	3.38e–14	0.14	3.38e–14	–3.25

Full Newton Iteration.

	$k=2,\ q_2=7$		$k=3,\ q_3=23$		$k=4,\ q_4=31$	
h	Error	Order	Error	Order	Error	Order
1.00e–1	1.31e–8		6.10e–8		6.10e–8	
5.00e–2	1.64e–10	6.31	2.43e–10	7.97	2.43e–10	7.97
2.50e–2	1.01e–12	7.34	1.08e–12	7.81	1.08e–12	7.81
1.25e–2	7.11e–15	7.14	3.55e–15	8.25	3.55e–15	8.25
6.25e–3	3.38e–14	–2.25	3.38e–14	–3.25	3.38e–14	–3.25

Table 4.2: Results for Example 4.2 based on the 4-stage Kuntzmann-Butcher formula of order $p=8$ using the associated collocation polynomial as the predictor, whence $q=r=4$.

Simple Iteration.

	$k=1,\ q_1=4$		$k=2,\ q_2=5$		$k=3,\ q_3=6$		$k=4,\ q_4=7$	
h	Error	Order	Error	Order	Error	Order	Error	Order
1.00e–1	4.70e–2	3.67	4.57e–3	4.66	3.95e–4	5.65	3.10e–5	6.64
5.00e–2	3.71e–3	3.83	1.81e–4	4.83	7.89e–6	5.82	3.12e–7	6.82
2.50e–2	2.60e–4	3.92	6.38e–6	4.91	1.39e–7	5.91	2.76e–9	6.90
1.25e–2	1.72e–5	3.96	2.12e–7	4.96	2.31e–9	5.96	2.31e–11	6.08
6.25e–3	1.11e–6		6.81e–9		3.73e–11		3.41e–13	

Modified Newton Iteration.

	$k=1,\ q_1=4$		$k=2,\ q_2=6$		$k=3,\ q_3=8$		$k=4,\ q_4=10$	
h	Error	Order	Error	Order	Error	Order	Error	Order
1.00e–1	4.70e–2	3.67	8.47e–5	5.60	3.50e–6	7.98	1.67e–6	8.69
5.00e–2	3.71e–3	3.83	1.74e–6	5.79	1.38e–8	7.97	4.06e–9	8.25
2.50e–2	2.60e–4	3.92	3.16e–8	5.89	5.51e–11	8.45	1.33e–11	7.66
1.25e–2	1.72e–5	3.96	5.34e–10	5.91	1.55e–13	–0.88	6.57e–14	0.45
6.25e–3	1.11e–6		8.85e–12		2.84e–13		4.80e–14	

Full Newton Iteration.

	$k=1,\ q_1=\bar{q}_1=4$		$k=2,\ q_2=7,\ \bar{q}_2=9$		$k=3,\ q_3=13,\ \bar{q}_3=19$	
h	Error	Order	Error	Order	Error	Order
1.00e–1	4.70e–2	3.67	1.64e–6	8.58	1.63e–6	8.65
5.00e–2	3.71e–3	3.83	4.29e–9	8.30	4.07e–9	8.25
2.50e–2	2.60e–4	3.92	1.36e–11	8.51	1.33e–11	7.71
1.25e–2	1.72e–5	3.96	3.73e–14	–2.35	6.39e–14	0.08
6.25e–3	1.11e–6		1.90e–13		6.04e–14	

Table 4.3: Results for Example 4.3 based on the 8^{th}-order extrapolation scheme shown in Figure 4.1 and using Heun's 3^{rd}-order explicit RK formula as predictor. Thus $p=8$, $q=3$ and $r=1$.

Full Newton Iteration, $k = 2$.

	Scheme A		Scheme B	
h	Error	Order	Error	Order
1.00e–1	2.60e–5		4.13e–6	
		6.99		4.49
5.00e–2	2.05e–7		1.83e–7	
		7.06		5.53
2.50e–2	1.54e–9		3.96e–9	
		7.06		5.80
1.25e–2	1.15e–11		7.10e–11	
		6.88		5.99
6.25e–3	6.04e–14		1.11e–12	

Table 4.4: Results for Schemes A and B described in Example 4.4. Both schemes are based on the 8^{th}-order extrapolation method, but Scheme A uses Heun's 2^{nd}-order formula only in its predictor and so satisfies (3.1.14), while Scheme B uses Heun's 2^{nd}- and 3^{rd}-order formulas in its predictor and does not satisfy (3.1.14). For both schemes, $p = 8$, $q = 2$, $r = 1$, $q_2 = 6$ and $\bar{q}_2 = 7$, but only Scheme A benefits from the higher order $\bar{q}_2 = 7$ attained by those methods which satisfy the assumptions of Theorem 3.3.4.

Simple Iteration.

	$k=1,\ p_1^H=6$		$k=2,\ p_2^H=4$		$k=3,\ p_3^H=3$		$k=4,\ p_4^H=3$	
h	Error	Order	Error	Order	Error	Order	Error	Order
1.00e–1	1.81e–6		3.37e–1		5.22e–1		1.66e–0	
		5.79		3.83		2.99		3.59
5.00e–2	3.29e–8		2.38e–2		6.56e–2		1.38e–1	
		5.90		3.94		3.02		3.43
2.50e–2	5.51e–10		1.54e–3		8.11e–3		1.28e–2	
		5.93		3.99		3.04		3.32
1.25e–2	9.03e–12		9.72e–5		9.90e–4		1.29e–3	
		8.73		4.00		3.03		3.20
6.25e–3	2.13e–14		6.07e–6		1.21e–4		1.40e–4	

Modified Newton Iteration.

	$k=1,\ p_1^R=6$		$k=2,\ p_2^R=3$		$k=3,\ p_3^R=3$		$k=4,\ p_4^R=3$	
h	Error	Order	Error	Order	Error	Order	Error	Order
1.00e–1	1.81e–6		1.75e–0		1.85e–0		1.85e–0	
		5.79		3.68		3.72		3.73
5.00e–2	3.29e–8		1.37e–1		1.40e–1		1.40e–1	
		5.90		3.45		3.47		3.47
2.50e–2	5.51e–10		1.26e–2		1.27e–2		1.27e–2	
		5.93		3.30		3.31		3.31
1.25e–2	9.03e–12		1.28e–3		1.28e–3		1.28e–3	
		8.73		3.19		3.20		3.20
6.25e–3	2.13e–14		1.39e–4		1.39e–4		1.39e–4	

Full Newton Iteration.

	$k=1,\ p_1^D=6$		$k=2,\ p_2^D=3$		$k=3,\ p_3^D=3$		$k=4,\ p_4^D=3$	
h	Error	Order	Error	Order	Error	Order	Error	Order
1.00e–1	1.81e–6		1.68e–0		1.85e–0		1.86e–0	
		5.78		3.63		3.73		3.73
5.00e–2	3.29e–8		1.36e–1		1.40e–1		1.40e–1	
		5.90		3.44		3.47		3.47
2.50e–2	5.51e–10		1.25e–2		1.27e–2		1.27e–2	
		5.93		3.30		3.31		3.31
1.25e–2	9.03e–12		1.27e–3		1.28e–3		1.28e–3	
		8.73		3.19		3.20		3.20
6.25e–3	2.13e–14		1.39e–4		1.39e–4		1.39e–4	

Table 4.5: Results for Example 4.5 based on the 3^{rd}-order DIRK formula shown in Figure 4.2 using Fehlberg's 8^{th}-order explicit RK formula as predictor. Thus, $p=3$, $q=8$ and $r=1$.

5 Concluding Remarks.

In this paper, we analyzed the order of iterated Runge-Kutta methods for three different iteration schemes. The results summarized in Theorem 1.1 show that on each iteration the order of the iterated RK method normally

- increases by one for fixed-point iteration,
- increases by two for modified Newton iteration,
- at least doubles for full Newton iteration if $C(q-1)$ is satisfied, where q is the order of the predictor, or almost doubles if $C(q-1)$ is not satisfied,

up to the maximum order of the underlying RK formula, although we have given examples in this paper and [11] that illustrate that exceptions to these rules occur.

We used B-series to prove Theorem 1.1 and associated results. The theoretical basis for all this work is the relation between the B-series for the internal stage values $Y_n^{(k)}$ on the k^{th} iteration and the B-series for the internal stage values $Y_n^{(k+1)}$ on the $k+1^{\text{st}}$ iteration. This relationship is best described in terms of the *growth functions* $H(t,q)$, $R(t,q)$ and $D(t,r,q)$, introduced in §2, that specify exactly which terms in the B-series of the iterated RK method equal the corresponding term of the underlying RK formula. Often only a few terms restrict the order of the iterated RK method. Partly because of this, the order of a particular iterated RK scheme might be higher than that guaranteed by Theorem 1.1, as explained in §3 and illustrated in §4 and [11].

In addition to forming the theoretical basis for iterated RK methods, this work sheds some light on several questions of practical importance, as illustrated by the following two examples of possibly problematic practices.

$\mathbb{P}1$. Most implicit RK codes do not force their scheme for computing the internal stage values to iterate more than once. This could lead to the RK method being of lower order than expected, and as a result possibly unreliable and/or inefficient. For example, suppose

- a code is based on the 4-stage Kuntzmann-Butcher formula [8, p. 205] of order 8,
- its error estimate is based on local extrapolation using the one-step-two-half-steps approach,
- the code uses the modified Newton iteration to compute the internal stage values, and
- it starts the iteration with predicted values computed from the associated collocation polynomial of degree 4 that interpolates y_{n-1} and $Y_{n-1,i}$, $i=1,\dots,4$, whence $q=4$.

If we perform just one modified Newton iteration (computing $Y_n^{(1)}$ from (1.13) and then using $Y_n^{(1)}$ in (1.14) to calculate $y_{n+1}^{(2)}$, as is usually done in practice), the iterated RK method will be of order 7 for some nonlinear problems. Worse still, its estimate of the error per step will be $O(h^9)$, because of the correlation between the one-step and two-half-step solutions. Thus, the error may be grossly underestimated. Moreover, 56 of the 286 error estimate coefficients of order 9 are zero, even though the corresponding true error coefficients are not. This may contribute to a further underestimation of the true error. On the other hand, if we perform two or more modified Newton iterations (computing $Y_n^{(1)}$ and $Y_n^{(2)}$ from (1.13) and then using $Y_n^{(2)}$ in (1.14) to calculate $y_{n+1}^{(3)}$), the iterated RK method will be of order 8, as expected, and in addition its principal error terms will agree with those of the underlying RK formula. Moreover, the error estimate will be asymptotically correct.

$\mathbb{P}2$. Another practice that may be problematic is using either $y_{n+1}^{(k)}$ or $y_{n+1}^{(k+1)}$ to advance the integration together with the difference

$$l_{n+1}^{(k)} = y_{n+1}^{(k+1)} - y_{n+1}^{(k)} = \sum_{t \in T} \left(\tilde{\phi}_{s+1}^{(k+1)}(t) - \tilde{\phi}_{s+1}^{(k)}(t) \right) F(t)(y_n) \frac{h^{\rho(t)}}{\rho(t)!}$$

as the basis for an error estimate, as suggested in [9] for example. Note that this technique is normally plausible only when $q_k < p$, where p is the order of the underlying RK formula, since otherwise $y_{n+1}^{(k)}$ and $y_{n+1}^{(k+1)}$ will typically be of the same order and so their difference will not provide an asymptotically correct estimate of the local error in $y_{n+1}^{(k)}$. But even if $q_k < p$, this technique may be problematic, since, although normally $y_{n+1}^{(k)}$ is of order q_k, $y_{n+1}^{(k+1)}$ is of order $q_{k+1} > q_k$, and $l_n^{(k)} = O(h^{q_k+1})$ as expected, many of the principal error terms of order $q_k + 1$ associated with $y_{n+1}^{(k)}$ may be zero, since, as shown in §3, $\tilde{\phi}_{s+1}^{(k)}(t) = \tilde{\phi}_{s+1}^{(k+1)}(t) = 1$ for all trees $t = [t_1, \ldots, t_m]$ satisfying $H(t_i, q) \leq k - 1$, $R(t_i, q) \leq k - 1$ or $D(t_i, r, q) \leq k - 1$ for $i = 1, \ldots, m$, as the case may be depending on which iteration scheme is used. For simple iteration with $q = 4$ and $k = 2$, for example, $\tilde{\phi}_{s+1}^{(2)}(t) = \tilde{\phi}_{s+1}^{(3)}(t) = 1$ for 39 of the 48 trees of order $q_2 + 1 = 7$, resulting in $y_{n+1}^{(2)}$ having a large portion — more than 81% — of its principal error terms of order $q_2 + 1 = 7$ equal to zero. Moreover, this example is typical of the general case. It has long been recognized [19, §7.5.3] as undesirable for the lower-order formula in an explicit RK formula pair to have principal error terms equal to zero and we believe it is undesirable for iterated RK methods too.

One weakness with our theory is that it is based on an asymptotic analysis which is valid only in the limit $h \to 0$. Hence it might not be relevant for stiff problems. The same objection is raised for the classical order analysis of RK formulas. To address this weakness, Kværnø [15] has performed a similar analysis for semi-explicit index 1 DAEs, a collection of problems which can be viewed as the limiting case of a large and important class of stiff ODEs.

Furthermore, we are not guaranteed that $Y^{(k)} \to Y$ as $k \to \infty$, except for sufficiently small h. This topic is carefully investigated by van Dorsselaer and Spijker [6].

References

[1] N. G. R. Broderick, C. M. de Sterke, and K. R. Jackson. Coupled mode equations with free carrier effects: A numerical solution. *J. Optical and Quantum Electronics*, 26:S219–S234, 1994.

[2] K. Burrage. The error behaviour of a general class of predictor-corrector methods. *Appl. Numer. Math.*, 8:201–216, 1991.

[3] K. Burrage, J. C. Butcher and F. H. Chipman. An implementation of singly-implicit Runge-Kutta methods. *BIT*, 20:326–340, 1980.

[4] J. C. Butcher. *The Numerical Analysis of Ordinary Differential Equations.* John Wiley & Sons, New York, 1987.

[5] C. M. de Sterke, K. R. Jackson, and B. D. Robert. Nonlinear coupled mode equations on a finite interval: A numerical procedure. *J. of the Optical Society of America B*, 8:403–412, 1991.

[6] J. L. M. van Dorsselaer and M. N. Spijker. The error committed by stopping Newton iterations in the numerical solution of stiff initial value problems. Technical Report TW–92–02, Dept. of Mathematics and Computer Science, Div. of Applied Mathematics, Leiden University, 1992.

[7] R. F. Enenkel. The implementation of parallel Runge-Kutta methods. Master's thesis, Computer Science Dept., University of Toronto, Toronto, Canada, 1988.

[8] E. Hairer, S. P. Nørsett and G. Wanner. *Solving Ordinary Differential Equations I, Nonstiff Problems.* Springer-Verlag, Berlin, 1987.

[9] P. J. van der Houwen, B. P. Sommeijer, and W. Couzy. Embedded diagonally implicit Runge-Kutta algorithms on parallel computers. *Math. Comp.*, 58:135–159, 1992.

[10] K. R. Jackson. A survey of parallel numerical methods for initial value problems for ordinary differential equations. *IEEE Trans. on Magnetics*, 27:3792–3797, 1991.

[11] K. R. Jackson, A. Kværnø and S. P. Nørsett. The use of Butcher series in the analysis of Newton-like iterations in Runge-Kutta formulas. To appear in *Applied Numerical Mathematics.*

[12] K. R. Jackson and S. P. Nørsett. The potential for parallelism in Runge-Kutta methods. Part 1: RK formulas in standard form. To appear in *SIAM J. Numer. Anal.*

[13] K. R. Jackson and S. P. Nørsett. The potential for parallelism in Runge-Kutta methods. Part 2: RK predictor-corrector formulas. In preparation.

[14] J. Kalvenes. Experimentation with parallel ODE-solvers. Technical Report NM–R8912, Mathematical Sciences Div., Norwegian Institute of Technology, Trondheim, Norway, 1986.

[15] A. Kværnø. The order of Runge-Kutta methods applied to semi-explicit DAEs of index 1, using Newton-type iterations to compute the internal stage values. Technical Report 2/1992, Mathematical Sciences Div., Norwegian Institute of Technology, Trondheim, Norway, 1992.

[16] S. P. Nørsett and H. H. Simonsen. Aspects of parallel Runge-Kutta methods. In A. Bellen, C. W. Gear and E. Russo, editors, *Numerical Methods for Ordinary Differential Equations*, pages 103–107, Berlin, 1989. Springer Verlag. Lecture Notes in Mathematics #1386, Proceedings of the l'Aquila Symposium 1987.

[17] S. P. Nørsett and P. G. Thomsen. User's guide for SIMPLE, a stiff system solver. Working notes.

[18] S. P. Nørsett and A. Wolfbrandt. Order conditions for Rosenbrock type methods. *Numer. Math.*, 32:1–15, 1979.

[19] L. F. Shampine. *Numerical Solution of Ordinary Differential Equations.* Chapman & Hall, New York, 1994.