©2026 Ada Academica https://adac.ee

Eur. J. Math. Anal. 6 (2026) 23
doi: 10.28924/ada/ma.6.23

A Second-Order Directional Correction Giving Genuine Eighth-Order Convergence for Nonlinear Systems with One Factorization per Iteration

Pairat Janngam, Chalermwut Comemuang*

Department of Mathematics, Faculty of Education, Buriram Rajabhat University, Buriram 31000, Thailand

pairat.jn@bru.ac.th; chalermwut.cm@bru.ac.th

** Correspondence: chalermwut.cm@bru.ac.th*

Abstract. Three-step methods that freeze the Jacobian at the current iterate and sample a second Jacobian at the Newton predictor are attractive for large systems, since one factorization serves the whole step. We show that such schemes cannot exceed order six. The obstruction is a single matrix, $R_k = \frac{1}{2}\mathcal{F}'(x_k)^{-1}\mathcal{F}''(x_k)[Q_k u_k, \cdot\,]$, which the two Jacobians cannot determine: we exhibit two problems agreeing in $\mathcal{F}(x_k)$, $\mathcal{F}'(x_k)$ and $\mathcal{F}'(y_k)$ whose values of R_k differ. For one equation $R_k = Q_k^2$, so the scalar theory is unaffected. We also determine when the barrier is inactive: if a constant starting vector or a set of affine equations confines the iteration to a one-dimensional set, the restricted problem is scalar and order eight is recovered. One further Jacobian, evaluated at a fourth-order predictor where the missing direction is exposed, supplies R_k and yields a method PC8 of genuine order eight. Written so that no weight matrix is assembled, PC8 costs one LU factorization plus $O(n^2)$ arithmetic per step, against about $4.7n^3$ for the usual implementation. Measured orders are 7.99–8.00 for PC8 and 5.97–6.00 for five published weights on the same problems.

1. Introduction

Solving $\mathcal{F}(x) = 0$ for $\mathcal{F} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is expensive mainly because of the linear algebra. Newton's method spends $\frac{2}{3}n^3$ operations on a factorization for every quadratic step [8, 11], and the natural way to buy a higher order cheaply is to keep that factorization and re-use it. The literature on multi-step variants is large and [9] surveys it; for a single equation the question of how much order a given number of evaluations can buy was settled by Kung and Traub [6], and the corresponding question for systems is what this paper is about. A separate line of work dispenses with the Jacobian and replaces it by divided differences [1]; such schemes trade arithmetic for evaluations and are complementary to what follows, since the obstruction

Received: 31 Aug 2026.

2020 *Mathematics Subject Classification.* 65H10; 65Y20; 41A25.

Key words and phrases. nonlinear systems; Newton-type methods; order of convergence; frozen Jacobian; computational efficiency.

identified here concerns which directional second derivatives an iteration can see, not how its first derivatives are obtained. This idea produces the three-step schemes

$$\begin{cases} y_k = x_k - \mathcal{F}'(x_k)^{-1}\mathcal{F}(x_k), \\ z_k = y_k - \bar{\tau}_k\, \mathcal{F}'(x_k)^{-1}\mathcal{F}(y_k), \\ x_{k+1} = z_k - \alpha_k\, \mathcal{F}'(x_k)^{-1}\mathcal{F}(z_k), \end{cases} \tag{1}$$

in which the weight matrices $\bar{\tau}_k, \alpha_k$ are assembled from $\mathcal{F}'(x_k)$ and one further Jacobian at y_k. Writing

$$t_k = \mathcal{F}'(x_k)^{-1}\mathcal{F}'(y_k), \qquad s_k = t_k^{-1}, \tag{2}$$

one takes $\bar{\tau}_k$ and α_k to be low-degree polynomials in t_k or s_k, chosen so that enough terms of the local error cancel. Schemes of this type appear in [2,3,10,13], and [14] organises them into a single family with sufficient conditions on $\bar{\tau}_k$ and α_k. The order usually claimed for them is eight.

How much order can a scheme of this shape attain? The question is worth asking because the answer is not the one the scalar theory suggests. It is six, and the reason is specific: there is one matrix that the two Jacobians cannot see, and both the middle and the outer step need it. Once that is understood the repair is short, and most of this paper is spent on the understanding rather than on the repair.

The local error of (1) is controlled by two matrices built from $\mathcal{F}''$. The first,

$$Q_k = \tfrac{1}{2}\mathcal{F}'(x_k)^{-1}\mathcal{F}''(x_k)[u_k, \cdot\,], \qquad u_k = \mathcal{F}'(x_k)^{-1}\mathcal{F}(x_k), \tag{3}$$

is standard and is visible in t_k. The second,

$$R_k = \tfrac{1}{2}\mathcal{F}'(x_k)^{-1}\mathcal{F}''(x_k)[Q_k u_k, \cdot\,], \tag{4}$$

is the same second derivative sampled in the direction $Q_k u_k$, which is of size $O(h^2)$ rather than $O(h)$. For a single equation the two are related by $R_k = Q_k^2$, because scalars commute and there is only one direction to sample in; this is why the corresponding scalar theory, for which raising the order is by now close to routine [4,7,12], gives no warning of what follows. For a system they are different matrices, and the difference is not small: on the boundary-value problem of Section 7 we measure $\|R_k - Q_k^2\|/\|\mathcal{F}(x_k)\|^2 \to 8.46$.

Section 4 shows that R_k is exactly what (1) needs and exactly what it cannot obtain. Because $y_k - x_k = -u_k$, the pair $\mathcal{F}'(x_k), \mathcal{F}'(y_k)$ reveals $\mathcal{F}''$ only in the direction u_k; the direction $Q_k u_k$ is invisible to it. Since the fifth-order accuracy of $\mathcal{F}(z_k)$ and the third-order accuracy of α_k both require R_k, the best available is one order less in each, and the composition gives order $4+2=6$. This is a structural limit rather than a property of any particular choice of weights: it holds for every $\bar{\tau}_k$ and α_k of the assumed form.

The remedy follows from the same analysis. A fourth-order predictor $\tilde{z}_k$ satisfies $\tilde{z}_k - y_k = -Q_k u_k + O(h^3)$, so one Jacobian evaluated there gives R_k to the accuracy required, and only through its action on vectors, hence in $O(n^2)$ operations. Section 5 states the corrected sufficient

conditions, in which the relation between the two weights turns out to be simply $\alpha_k = \bar{\tau}_k + R_k$, and the method PC8 that realises them. Section 6 treats the implementation and the cost, and Section 7 the numerical evidence.

Section 4 also determines when the barrier is inactive. If the iteration stays inside a one-dimensional subset of $\mathbb{R}^n$, the restricted problem is a single equation, $R_k = Q_k^2$ holds along the whole orbit, and order eight is attained. This happens in two ways that are easy to trigger without noticing: when the starting vector and the solution are both constant for a system of identical equations, and when enough of the equations are affine. Both are common in test collections, so the distinction matters whenever an order is measured numerically.

2. Notation

Let $x^* \in D$ be a simple zero of $\mathcal{F}$, so that $\mathcal{F}'(x^*)$ is nonsingular, and let $\mathcal{F}$ be four times continuously differentiable near x^*. For the current iterate x_k put

$$J_k = \mathcal{F}'(x_k), \qquad u_k = J_k^{-1}\mathcal{F}(x_k), \qquad h = \mathcal{F}(x_k), \tag{5}$$

and let $O(h^m)$ abbreviate $O(\|h\|^m)$; thus $u_k = O(h^1)$. We regard $\mathcal{F}''(x)$ as the symmetric bilinear map $(a,b) \mapsto \mathcal{F}''(x)[a,b] \in \mathbb{R}^n$ and, for a fixed a, write $\mathcal{F}''(x)[a,\cdot\,]$ for the matrix $v \mapsto \mathcal{F}''(x)[a,v]$; the third derivative is treated in the same way. With this convention we set

$$Q_k = \tfrac{1}{2}J_k^{-1}\mathcal{F}''(x_k)[u_k,\cdot\,], \qquad \bar{d}_k = -\tfrac{1}{6}J_k^{-1}\mathcal{F}'''(x_k)[u_k,u_k,\cdot\,], \tag{6}$$

$$R_k = \tfrac{1}{2}J_k^{-1}\mathcal{F}''(x_k)[Q_k u_k,\cdot\,], \tag{7}$$

so that $Q_k = O(h^1)$ and $\bar{d}_k = R_k = O(h^2)$. Note that $Q_k u_k = \tfrac{1}{2}J_k^{-1}\mathcal{F}''(x_k)[u_k,u_k]$, the leading term of the Newton error.

Two expansions are used constantly. Taylor expansion at x_k, together with $y_k - x_k = -u_k$ and $\mathcal{F}(x_k) = J_k u_k$, gives

$$v_k := J_k^{-1}\mathcal{F}(y_k) = Q_k u_k + \bar{d}_k u_k + O(h^4) = O(h^2) \tag{8}$$

and

$$t_k = I - 2Q_k - 3\bar{d}_k + O(h^3), \qquad P_k := I - t_k = O(h^1). \tag{9}$$

By the Neumann series,

$$s_k = t_k^{-1} = I + P_k + P_k^2 + O(h^3) = I + 2Q_k + 3\bar{d}_k + 4Q_k^2 + O(h^3). \tag{10}$$

Finally, we say that $\{x_k\}$ converges with order p if $\|\mathcal{F}(x_{k+1})\| = O(\|\mathcal{F}(x_k)\|^p)$ as $k \to \infty$; for a simple zero this agrees with the usual definition in terms of $\|x_k - x^*\|$.

Table 1. The defect (12) on the test systems of Section 7 with $n = 20$. A scalar equation is included for comparison.

system	$t = 10^{-2}$	$t = 10^{-3}$	$t = 10^{-4}$
$x^3 - x - 8$ (scalar)	0	0	$2.42 \cdot 10^{-19}$
S1, trigonometric	$1.2436 \cdot 10^{-1}$	$1.2229 \cdot 10^{-1}$	$1.2209 \cdot 10^{-1}$
S2, cyclic quadratic	$2.2026 \cdot 10^{-2}$	$2.1605 \cdot 10^{-2}$	$2.1564 \cdot 10^{-2}$
S3, boundary-value problem	8.9717	8.4927	8.4585

3. The two second-order matrices

The whole analysis rests on the difference between Q_k^2 and R_k, so we record it first.

Proposition 3.1. *For every $v \in \mathbb{R}^n$,*

$$Q_k^2 v = \tfrac{1}{4} J_k^{-1} \mathcal{F}''[u_k,\, J_k^{-1}\mathcal{F}''[u_k, v]], \qquad R_k v = \tfrac{1}{4} J_k^{-1} \mathcal{F}''[J_k^{-1}\mathcal{F}''[u_k, u_k],\, v], \tag{11}$$

where $\mathcal{F}'' = \mathcal{F}''(x_k)$. If $n = 1$ these agree identically. If $n \geq 2$ they differ for generic $\mathcal{F}$.

Proof. Both formulas follow from (6)–(7) by substitution. For $n = 1$ every quantity is a scalar and both sides reduce to $\tfrac{1}{4}(\mathcal{F}''/J_k)^2 u_k^2 v$. For $n \geq 2$ the two expressions contract the bilinear map $\mathcal{F}''$ in different arguments: in $Q_k^2 v$ the vector u_k occupies the first slot of the outer $\mathcal{F}''$ and v enters only through the inner one, while in $R_k v$ the outer $\mathcal{F}''$ receives $\mathcal{F}''[u_k, u_k]$ and v directly. Symmetry of $\mathcal{F}''$ in its two arguments does not exchange the two patterns: in Q_k^2 the tensor $\mathcal{F}''$ is contracted with u_k in one slot of each of its two occurrences, while in R_k the inner occurrence is contracted with u_k in both slots and the outer one is left free. The two contractions agree only when every index collapses, that is when $n = 1$. A two-dimensional example is $J_k = I$, $\mathcal{F}''[a, b] = (a_1 b_1,\, a_1 b_2 + a_2 b_1)^\top$, $u_k = (1, 1)^\top$, for which $Q_k^2 \neq R_k$. □

Proposition 3.1 also delimits the scope of everything that follows. For a single equation the identity $R_k = Q_k^2$ is available, and the constructions that exploit it are legitimate: the auxiliary-function and modified homotopy-perturbation schemes of [4], the linear-combination weights of [7], and the general algorithm of [12] for producing Newton formulae of arbitrary order all remain valid as stated. The difficulty arises only on passing to $\mathbb{R}^n$, and it arises immediately.

The distinction is not a curiosity of pathological examples. Table 1 evaluates

$$\delta(x) = \frac{\|R_k - Q_k^2\|_F}{\|\mathcal{F}(x)\|^2} \tag{12}$$

along $x = x^* + t\,d$, $t \to 0$, for a fixed random unit vector d, on the three systems used in Section 7. A limit different from zero means that $R_k - Q_k^2$ is itself of size $O(h^2)$, that is, as large as either term.

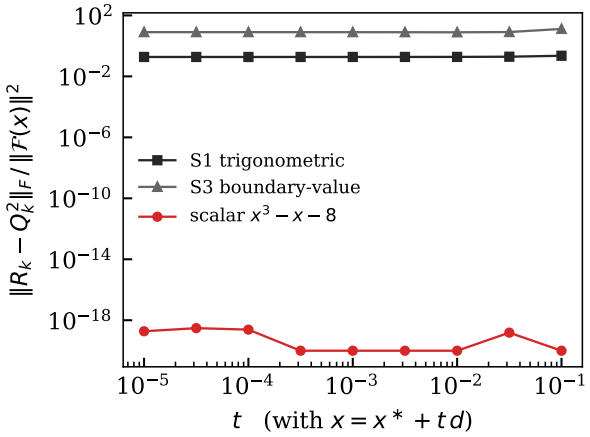


Figure 1. The defect (12) as $t \to 0$. For a single equation it falls to the level of rounding; for the two systems it settles at a nonzero constant, so $R_k - Q_k^2$ is of the same order as either term separately.

Figure 1 plots the same quantity over a range of t: the scalar curve descends into rounding error while the two system curves flatten out, which is the picture the table records at three points.

4. What two Jacobians determine

We now derive the exact local expansions for (1). Nothing beyond Taylor's theorem is used.

Lemma 4.1. *Let $z_k = y_k - \bar{\tau}_k J_k^{-1}\mathcal{F}(y_k)$ with $\bar{\tau}_k = I + O(h^1)$. Then*

$$\mathcal{F}(z_k) = J_k(I - t_k\bar{\tau}_k + R_k)v_k + O(h^5). \tag{13}$$

Proof. Expand $\mathcal{F}$ at y_k. Using $\mathcal{F}(y_k) = J_k v_k$, $\mathcal{F}'(y_k) = J_k t_k$ and $z_k - y_k = -\bar{\tau}_k v_k = -v_k + O(h^3)$,

$$\mathcal{F}(z_k) = J_k v_k - J_k t_k \bar{\tau}_k v_k + \tfrac{1}{2}\mathcal{F}''(x_k)[v_k, v_k] + O(h^5),$$

the replacement of $\mathcal{F}''(y_k)$ by $\mathcal{F}''(x_k)$ costing $O(h^5)$ because $v_k = O(h^2)$. By (8), $v_k = Q_k u_k + O(h^3)$, so $\frac{1}{2}J_k^{-1}\mathcal{F}''(x_k)[v_k, v_k] = R_k v_k + O(h^5)$. □

Lemma 4.2. *Under the hypotheses of Lemma 4.1,*

$$J_k^{-1}\mathcal{F}'(z_k) = I - P_k - 2R_k + O(h^3), \qquad \left(J_k^{-1}\mathcal{F}'(z_k)\right)^{-1} = I + P_k + P_k^2 + 2R_k + O(h^3). \tag{14}$$

Proof. From (8) and $\bar{\tau}_k = I + O(h^1)$ we get $z_k - x_k = -(I + Q_k)u_k + O(h^3)$. Expanding $\mathcal{F}'$ at x_k and using (6)–(7),

$$J_k^{-1}\mathcal{F}'(z_k) = I - 2Q_k - 2R_k - 3\bar{d}_k + O(h^3),$$

which is the first identity by (9). The second follows from the Neumann series and $P_k^2 = 4Q_k^2 + O(h^3)$. □

Both lemmas single out R_k. It is therefore worth asking what the data available to (1) actually determine.

Lemma 4.3. *Let $n \geq 2$. There exist two maps $\mathcal{F}, \tilde{\mathcal{F}} : \mathbb{R}^n \to \mathbb{R}^n$ and a point x_k at which*

$$\mathcal{F}(x_k) = \tilde{\mathcal{F}}(x_k), \qquad \mathcal{F}'(x_k) = \tilde{\mathcal{F}}'(x_k), \qquad \mathcal{F}'(y_k) = \tilde{\mathcal{F}}'(y_k), \tag{15}$$

so that y_k, t_k, Q_k, $\bar{d}_k$ and the vector $Q_k u_k$ are common to both, while $R_k \neq \tilde{R}_k$. Consequently there is no map W for which $R_k = W(\mathcal{F}'(x_k), \mathcal{F}'(y_k)) + O(h^3)$ holds for every sufficiently differentiable $\mathcal{F}$.

Proof. Take $n = 2$, $x_k = 0$, and

$$\mathcal{F}(x) = a + Jx + \tfrac{1}{2}B[x, x], \qquad a = e_1, \quad J = I,$$

where B is the symmetric bilinear map with $B[e_1, e_1] = 2e_2$ and every other slot zero. Then $u_k = J^{-1}\mathcal{F}(0) = e_1$, $y_k = -e_1$, and $Q_k u_k = \tfrac{1}{2}B[u_k, u_k] = e_2$.

Let D be the symmetric bilinear map with $D[e_2, e_2] = e_1$ and every other slot zero, so that $D[e_1, \cdot\,] = 0$, and put $\tilde{\mathcal{F}}(x) = a + Jx + \tfrac{1}{2}(B + D)[x, x]$. Because $D[e_1, \cdot\,] = 0$ we have $\tilde{\mathcal{F}}(0) = a$, $\tilde{\mathcal{F}}'(0) = J$ and

$$\tilde{\mathcal{F}}'(y_k) = J - (B + D)[e_1, \cdot\,] = J - B[e_1, \cdot\,] = \mathcal{F}'(y_k),$$

which is (15); likewise $\tilde{Q}_k = Q_k$ and $\tilde{Q}_k u_k = Q_k u_k = e_2$, and $\bar{d}_k = \tilde{\bar{d}}_k = 0$ since both maps are quadratic. However

$$\tilde{R}_k - R_k = \tfrac{1}{2}J^{-1}D[e_2, \cdot\,] \neq 0,$$

of Frobenius norm $\tfrac{1}{2}$. A function of the data (15) must take the same value on $\mathcal{F}$ and $\tilde{\mathcal{F}}$, so it cannot equal R_k on both, and the error cannot be $O(h^3)$ because the discrepancy is independent of h.

For $n > 2$ append affine components: replace $\mathcal{F}$ by $\hat{\mathcal{F}}(x) = (\mathcal{F}(x_1, x_2), x_3, \ldots, x_n)$ and $\tilde{\mathcal{F}}$ by $\hat{\tilde{\mathcal{F}}}$ in the same way. The added components contribute nothing to $\mathcal{F}''$, the Jacobian becomes block diagonal with an identity block, and u_k, Q_k, R_k keep their two-dimensional blocks and vanish elsewhere. The two maps again share the data (15) while their values of R_k still differ by $\tfrac{1}{2}$. □

The construction isolates the mechanism. The pair of Jacobians determines the slice of $\mathcal{F}''(x_k)$ in the direction u_k, since $y_k - x_k = -u_k$ is the only displacement at which the second Jacobian is sampled. The matrix R_k is the slice in the direction $Q_k u_k$, and the two slices can be varied independently whenever $Q_k u_k$ is not parallel to u_k, which is the generic situation once $n \geq 2$.

Since $y_k - x_k = -u_k$, the second Jacobian probes $\mathcal{F}''$ only along u_k. The matrix R_k probes it along $Q_k u_k$, and by Proposition 3.1 this is genuinely new information once $n \geq 2$. Combining the three lemmas gives the main structural result.

Theorem 4.4 (Order barrier)**.** *Suppose $\bar{\tau}_k$ and α_k in (1) are matrix-valued functions of $\mathcal{F}'(x_k)$ and $\mathcal{F}'(y_k)$. If $R_k \neq Q_k^2 + O(h^3)$, then the order of convergence of (1) is at most six. The bound is sharp.*

Proof. The argument has four steps.

Step 1: what the middle step needs. By Lemma 4.1, $\mathcal{F}(z_k) = O(h^5)$ holds if and only if $I - t_k\bar{\tau}_k + R_k = O(h^3)$, that is if and only if

$$\bar{\tau}_k = t_k^{-1}(I + R_k) + O(h^3). \tag{16}$$

Every ingredient on the right except R_k is a function of $\mathcal{F}'(x_k)$ and $\mathcal{F}'(y_k)$, so (16) asks $\bar{\tau}_k$ to supply R_k and nothing else.

Step 2: the deficiency is of size exactly $O(h^2)$. Put $\Delta_k = \bar{\tau}_k - t_k^{-1}(I + R_k)$. By hypothesis $\bar{\tau}_k$ is a function of $\mathcal{F}'(x_k)$ and $\mathcal{F}'(y_k)$, and by Lemma 4.3 no such function agrees with $t_k^{-1}(I + R_k)$ to within $O(h^3)$ for every $\mathcal{F}$: on the pair constructed there all the admissible data coincide while the two values of R_k differ by a fixed amount, so a single function cannot match both. Hence $\Delta_k \neq O(h^3)$. Conversely $\Delta_k = O(h^2)$ is attainable, since $\bar{\tau}_k = t_k^{-1}(I+Q_k^2)$ is admissible — Q_k^2 is a polynomial in t_k by (9) — and leaves $\Delta_k = t_k^{-1}(Q_k^2 - R_k)$. That this is $O(h^2)$ needs no cancellation between the two terms: $Q_k = O(h^1)$ by (6), so $Q_k^2 = O(h^2)$; and $R_k = \frac{1}{2}J_k^{-1}\mathcal{F}''(x_k)[Q_k u_k, \cdot\,]$ is linear in its first argument, which is $Q_k u_k = O(h^2)$, so $R_k = O(h^2)$ as well. Their difference is therefore $O(h^2)$, and $t_k^{-1} = I + O(h^1)$ does not change the order. That it is no *smaller* than $O(h^2)$ is exactly the hypothesis of the theorem, $R_k \neq Q_k^2 + O(h^3)$; Proposition 3.1 and Lemma 4.3 show that this hypothesis is not vacuous, since it fails only on the set where the two contractions of $\mathcal{F}''$ happen to agree. The shortfall therefore sits exactly at order h^2, and no choice of $\bar{\tau}_k$ can remove it or make it worse.

Step 3: one order is lost in the middle step. From (13) and Step 2,

$$\mathcal{F}(z_k) = J_k(I - t_k\bar{\tau}_k + R_k)v_k + O(h^5) = -J_k t_k \Delta_k v_k + O(h^5),$$

and since $v_k = O(h^2)$ by (8),

$$\mathcal{F}(z_k) = O(h^4) \tag{17}$$

instead of the $O(h^5)$ that (16) would have produced.

Step 4: a second order is lost in the outer step. Expand $\mathcal{F}$ at z_k, and write $E_k = \left(J_k^{-1}\mathcal{F}'(z_k)\right)^{-1} - \alpha_k$; then

$$\mathcal{F}(x_{k+1}) = \mathcal{F}'(z_k)\, E_k\, J_k^{-1}\mathcal{F}(z_k) + O(\|\mathcal{F}(z_k)\|^2). \tag{18}$$

By Lemma 4.2 the requirement is $\alpha_k = I + P_k + P_k^2 + 2R_k + O(h^3)$, which asks for R_k again; the reasoning of Step 2 applies verbatim and gives $E_k \neq O(h^3)$ with $E_k = O(h^2)$ attainable, the admissible choice being $\alpha_k = I + P_k + P_k^2 + 2Q_k^2$. Combining with (17), the first term of (18) is $O(h^2)\cdot O(h^4) = O(h^6)$ and the second is $O(h^8)$, so $\mathcal{F}(x_{k+1}) = O(h^6)$.

The two losses are independent. The first occurs because $\bar{\tau}_k$ cannot see R_k, the second because α_k cannot, and no choice of one repairs the other, both being functions of the same pair of Jacobians. That is why the shortfall from eight is two orders rather than one.

Sharpness: taking $\bar{\tau}_k = I + P_k + P_k^2$ and $\alpha_k = I + P_k + 2P_k^2$ attains $O(h^2)$ in both places, and Table 3 below reports measured order 6.00 for this and for four weights from the literature. □

Remark 4.5. Condition $R_k = Q_k^2 + O(h^3)$ appears in the analysis of schemes of the form (1); in [14] it is stated explicitly as a hypothesis, in the equivalent form

$$\tfrac{1}{2} J_k^{-1} \mathcal{F}''(x_k)\, Q_k\, J_k^{-1} \mathcal{F}(x_k) = Q_k^2 + O(h^3). \tag{19}$$

Theorem 4.4 concerns exactly the complementary case. Under (19) the barrier does not apply and order eight is available; Section 4.1 describes the problems for which (19) holds.

4.1. **When the barrier is inactive.** Theorem 4.4 says something about $\mathcal{F}$ as well as about the scheme, since its hypothesis fails whenever $R_k = Q_k^2 + O(h^3)$. It is worth knowing which problems fall into that case, and two of them are easy to construct by accident rather than on purpose.

Proposition 4.6. *Let $e = (1, \dots, 1)^\top$ and $V = \operatorname{span}\{e\}$. If $\mathcal{F}(ce) \in V$ and $\mathcal{F}'(ce)V \subseteq V$ for every c near the solution, then V is invariant under all three steps of (1); if in addition $x_0 \in V$ and $x^* \in V$, the iteration is that of a single equation and (19) holds along the whole orbit.*

Proof. Under the hypotheses $\mathcal{F}(x) \in V$ and $\mathcal{F}'(x)^{-1}V = V$ for $x \in V$, so each of u_k, v_k and $J_k^{-1}\mathcal{F}(z_k)$ lies in V, and the weights map V to V because t_k does. Restricted to the line V the system is one scalar equation, and Proposition 3.1 applies. □

The hypotheses are met by systems built from one repeated formula. For the trigonometric system S1 of Section 7, $x = ce$ gives $2x_i - \sum_{j \le 4} x_j = -2c$ for every i, so $\mathcal{F}(x) = (c - \cos 2c)e$, and $\mathcal{F}'(x) = I + \sin(-2c)G$ with $Ge = -2e$ preserves V; the cyclic system S2 behaves the same way. Both have constant solutions, so a constant starting vector — the natural choice for such a system — produces a scalar orbit and order eight. Table 4 shows what happens when the starting vector is moved off V: the measured order falls from 8.12 to 6.13 without anything in the method changing.

Proposition 4.7. *Suppose some component $\mathcal{F}_i$ of $\mathcal{F}$ is affine, $\mathcal{F}_i(x) = a^\top x - b$, and let the weights of (1) be polynomials p, q in t_k with $p(1) = q(1) = 1$. Then $\mathcal{F}_i(y_k) = \mathcal{F}_i(z_k) = \mathcal{F}_i(x_{k+1}) = 0$ exactly, for every k and every starting point.*

Proof. Row i of J_k is $a^\top$ and does not depend on the point, so $e_i^\top J_k = a^\top$ and $e_i^\top \mathcal{F}'(y_k) = a^\top$ as well. Since $\mathcal{F}_i$ is affine, $\mathcal{F}_i(y_k) = \mathcal{F}_i(x_k) + a^\top(y_k - x_k) = \mathcal{F}_i(x_k) - e_i^\top J_k J_k^{-1}\mathcal{F}(x_k) = 0$. Next, $e_i^\top J_k t_k = e_i^\top \mathcal{F}'(y_k) = a^\top = e_i^\top J_k$, so the covector $e_i^\top J_k$ is a left eigenvector of t_k with eigenvalue 1 and hence $e_i^\top J_k\, p(t_k) = p(1)\, e_i^\top J_k = e_i^\top J_k$. Therefore

$$\mathcal{F}_i(z_k) = \mathcal{F}_i(y_k) - e_i^\top J_k\, p(t_k) J_k^{-1}\mathcal{F}(y_k) = -e_i^\top \mathcal{F}(y_k) = -\mathcal{F}_i(y_k) = 0,$$

and the same computation with q gives $\mathcal{F}_i(x_{k+1}) = 0$. □

The normalisation $p(1) = q(1) = 1$ is not a restriction: it is forced by $\bar{\tau}_k = I + O(h^1)$, so every consistent weight satisfies it. The consequence is that a system containing r affine equations confines the iteration to an affine subspace of dimension $n-r$ from the first step onwards. When $n - r = 1$ the restricted problem is a single equation and, by Proposition 3.1, assumption (19) holds along the whole orbit.

This is worth checking against published test data. Example 1 of [10] is a 2×2 system whose second equation is $x_1 + x_2 - 1 = 0$; Example 2 is a 3×3 system with two affine equations. In both cases $n - r = 1$. Measuring the local order of NLM8 on Example 1 of [10] along the direction the orbit is confined to returns 7.994, while along a generic direction it returns 5.999; the remaining four test systems of [10] are started from constant vectors and fall under Proposition 4.6.

Proposition 4.8. *Let the weights of (1) be those of the scheme of Wang [13], obtained there by an undetermined-parameter technique with $M_k = J_k^{-1}(J_k - \mathcal{F}'(y_k))$ and the constants fixed at $g_1 = \frac{5}{4}$, $g_2 = \frac{3}{2}$. Then*

$$\bar{\tau}_k = I + P_k + P_k^2 + Q_k^2 + O(h^3), \qquad \alpha_k = I + P_k + P_k^2 + 2Q_k^2 + O(h^3), \tag{20}$$

that is, exactly the conditions (22)–(23) of Theorem 5.1 with R_k replaced by Q_k^2. Consequently that scheme has order six.

Proof. By definition $M_k = J_k^{-1}(J_k - \mathcal{F}'(y_k)) = I - t_k = P_k$, so the weights read

$$\bar{\tau}_k = I + (I + \tfrac{5}{4}P_k)P_k = I + P_k + \tfrac{5}{4}P_k^2, \qquad \alpha_k = I + (I + \tfrac{3}{2}P_k)P_k = I + P_k + \tfrac{3}{2}P_k^2. \tag{21}$$

Since $P_k = I - t_k = 2Q_k + 3\bar{d}_k + O(h^3)$ by (9), and $\bar{d}_k = O(h^2)$, we have $P_k^2 = 4Q_k^2 + O(h^3)$, whence $\frac{5}{4}P_k^2 = P_k^2 + Q_k^2 + O(h^3)$ and $\frac{3}{2}P_k^2 = P_k^2 + 2Q_k^2 + O(h^3)$, which is (20). The order then follows from Theorem 4.4, the weights being polynomials in t_k. □

Proposition 4.8 illustrates how little freedom the shape (1) leaves. The weights of [13] are obtained by fitting two scalar constants rather than by imposing (22), and they land on the only values consistent with $R_k = Q_k^2$. Two constructions with no intermediate step in common therefore produce the same pair of weights, which is what one expects of a limitation belonging to the class rather than to any particular derivation.

5. Corrected conditions and the method PC8

Removing (19) from the analysis costs nothing and yields conditions that are, if anything, tidier than the ones they replace.

Theorem 5.1. *In scheme (1):*

(i) *$\mathcal{F}(z_k) = O(h^5)$ if and only if $\bar{\tau}_k = t_k^{-1}(I + R_k) + O(h^3)$, equivalently*

$$\bar{\tau}_k = I + P_k + P_k^2 + R_k + O(h^3); \tag{22}$$

(ii) *if* (i) *holds, the order is eight if and only if*

$$\alpha_k = I + P_k + P_k^2 + 2R_k + O(h^3), \tag{23}$$

that is, if and only if $\alpha_k = \bar{\tau}_k + R_k + O(h^3)$.

Proof. Part (i) is Lemma 4.1 combined with (10). For (ii), $\mathcal{F}(z_k) = O(h^5)$ makes the remainder in (18) equal to $O(h^{10})$, so $\mathcal{F}(x_{k+1}) = O(h^8)$ exactly when $E_k = O(h^3)$, which by Lemma 4.2 is (23). □

Replacing R_k by Q_k^2 in (22)–(23) returns the conditions found in the literature, which is a useful consistency check and also shows that no other discrepancy is present: the entire difference between order six and order eight sits in one matrix.

5.1. **Obtaining R_k.** Let

$$\tilde{z}_k = y_k - (2I - t_k)J_k^{-1}\mathcal{F}(y_k) \tag{24}$$

be the fourth-order two-step iterate. From (9) and (8), $(2I - t_k)v_k = v_k + O(h^3)$ and hence

$$\tilde{z}_k - y_k = -Q_k u_k + O(h^3), \tag{25}$$

which is the direction we need. Expanding $\mathcal{F}'$ at y_k,

$$J_k^{-1}\mathcal{F}'(\tilde{z}_k) = t_k - 2R_k + O(h^3), \tag{26}$$

so that

$$R_k = \tfrac{1}{2}(t_k - J_k^{-1}\mathcal{F}'(\tilde{z}_k)) + O(h^3). \tag{27}$$

The error in (27) is exactly the tolerance allowed by Theorem 5.1. What matters for the implementation is that R_k is needed only through its action on a vector,

$$R_k w = \tfrac{1}{2}J_k^{-1}(\mathcal{F}'(y_k)w - \mathcal{F}'(\tilde{z}_k)w), \tag{28}$$

which costs two matrix–vector products and one pair of triangular solves.

Method PC8. Given x_k, compute x_{k+1} as follows.

(1) Evaluate $\mathcal{F}(x_k)$ and $J_k = \mathcal{F}'(x_k)$, and factor $J_k = LU$. Set $u_k = J_k^{-1}\mathcal{F}(x_k)$, $y_k = x_k - u_k$.
(2) Evaluate $\mathcal{F}(y_k)$ and $\mathcal{F}'(y_k)$. Set $v_k = J_k^{-1}\mathcal{F}(y_k)$ and define

$$T(w) = J_k^{-1}(\mathcal{F}'(y_k)w), \qquad P(w) = w - T(w).$$

(3) Predictor: $\tilde{z}_k = y_k - (2v_k - T(v_k))$.
(4) Evaluate $\mathcal{F}'(\tilde{z}_k)$ and define $R(\cdot)$ by (28).
(5) $z_k = y_k - [v_k + P(v_k) + P(P(v_k)) + R(v_k)]$.
(6) Evaluate $\mathcal{F}(z_k)$, set $q_k = J_k^{-1}\mathcal{F}(z_k)$ and

$$x_{k+1} = z_k - [q_k + P(q_k) + P(P(q_k)) + 2R(q_k)].$$

Theorem 5.2. *Method PC8 converges locally with order eight.*

Proof. We verify that the two conditions of Theorem 5.1 hold for the operators actually formed in steps 5 and 6.

First, the predictor. By (9), $2I - t_k = I + 2Q_k + 3\bar{d}_k + O(h^3)$, and by (8), $v_k = Q_k u_k + \bar{d}_k u_k + O(h^4)$; since $Q_k v_k = O(h^3)$ and $\bar{d}_k u_k = O(h^3)$,

$$\tilde{z}_k - y_k = -(2I - t_k)v_k = -v_k + O(h^3) = -Q_k u_k + O(h^3),$$

which is step 3 of the method. Expanding $\mathcal{F}'$ at y_k along this displacement,

$$J_k^{-1}\mathcal{F}'(\tilde{z}_k) = t_k + J_k^{-1}\mathcal{F}''(x_k)[\tilde{z}_k - y_k, \cdot\,] + O(h^3) = t_k - 2R_k + O(h^3),$$

so the operator $R(\cdot)$ of step 4, namely $R(w) = \frac{1}{2}J_k^{-1}(\mathcal{F}'(y_k)w - \mathcal{F}'(\tilde{z}_k)w)$, satisfies $R(w) = (R_k + O(h^3))w$. Write $\hat{R}_k = R_k + O(h^3)$ for the matrix it represents.

Step 5 forms $z_k = y_k - \hat{\bar{\tau}}_k v_k$, where by construction $\hat{\bar{\tau}}_k = I + P_k + P_k^2 + \hat{R}_k$. Substituting $\hat{R}_k = R_k + O(h^3)$,

$$\hat{\bar{\tau}}_k = I + P_k + P_k^2 + (R_k + O(h^3)) = (I + P_k + P_k^2 + R_k) + O(h^3), \tag{29}$$

and the right-hand side is condition (22) verbatim, the $O(h^3)$ being precisely the tolerance that condition allows. Theorem 5.1(i) therefore gives $\mathcal{F}(z_k) = O(h^5)$.

Step 6 forms $x_{k+1} = z_k - \hat{\alpha}_k q_k$ with $\hat{\alpha}_k = I + P_k + P_k^2 + 2\hat{R}_k$, and the same substitution gives

$$\hat{\alpha}_k = I + P_k + P_k^2 + 2(R_k + O(h^3)) = (I + P_k + P_k^2 + 2R_k) + O(h^3), \tag{30}$$

which is condition (23). Since (29) has already secured $\mathcal{F}(z_k) = O(h^5)$, Theorem 5.1(ii) applies and the order is eight.

The $O(h^3)$ errors carried by $\hat{R}_k$ are harmless at both places: in step 5 the perturbation multiplies $v_k = O(h^2)$ and contributes $O(h^5)$ to $\mathcal{F}(z_k)$, which is the order already claimed; in step 6 it multiplies $q_k = J_k^{-1}\mathcal{F}(z_k) = O(h^5)$ and contributes $O(h^8)$. □

By Theorem 4.4 the third Jacobian is not a convenience but a necessity, and since ZMO1 and ZMO3 of [14] already evaluate three while attaining order six, what distinguishes PC8 is where the third evaluation is made rather than that it is made at all.

6. Implementation and cost

Two points determine the cost of a step, and both concern the linear algebra rather than the evaluations.

The first is that the weight matrices should never be formed. In (1) the matrices $\bar{\tau}_k$ and α_k are applied to the single vectors $J_k^{-1}\mathcal{F}(y_k)$ and $J_k^{-1}\mathcal{F}(z_k)$, so a polynomial weight can be evaluated by Horner's rule on those vectors:

$$(a_0 I + a_1 t_k + a_2 t_k^2)w = a_0 w + a_1 T(w) + a_2 T(T(w)), \qquad T(w) = J_k^{-1}(\mathcal{F}'(y_k)w). \tag{31}$$

Each application of T is one matrix–vector product and one pair of triangular solves, $3n^2$ operations in total. The iterates are unchanged: in our double-precision tests the largest relative discrepancy between the two implementations over $n \le 2000$ was $1.8 \cdot 10^{-16}$.

Table 2. Leading cost of one iteration. Evaluations are counted in scalar component evaluations, as is customary.

scheme	**order**	**evaluations**	LU	**leading arithmetic**
two Jacobians, formed	6	$2n^2+3n$	1	$\approx 4.7\,n^3$
two Jacobians, operator form	6	$2n^2+3n$	1	$\frac{2}{3}n^3+O(n^2)$
three Jacobians, formed	6	$3n^2+3n$	1	$\approx 6.3\,n^3$
PC8	8	$3n^2+3n$	1	$\frac{2}{3}n^3+O(n^2)$

The second point is that forming the matrix t_k is expensive and easy to overlook. It is the solution of a linear system with n right-hand sides, $2n^3$ operations beyond the factorization, and each additional power t_k^2 is a further matrix–matrix product of $2n^3$. A literal implementation of a weight of degree two therefore costs

$$\underbrace{\tfrac{2}{3}n^3}_{\text{factorization}} + \underbrace{2n^3}_{t_k} + \underbrace{2n^3}_{t_k^2} \approx 4.7\,n^3, \tag{32}$$

seven times the factorization itself. Cost comparisons that count factorizations, triangular solves and matrix–vector products but not matrix–matrix products will therefore rank such schemes far too favourably. Written as in (31), the same scheme costs $\frac{2}{3}n^3+O(n^2)$.

The distinction matters even more when $\mathcal{F}'$ is sparse. The matrix t_k is in general dense even when J_k is not, so forming it destroys the sparsity and costs $O(n^2)$ storage. On a tridiagonal system of order 8000 a sparse factorization took 3.1 ms and a single sparse solve 0.20 ms on our machine, so assembling t_k by n solves would cost about 1.6 s, some five hundred times the factorization, before any matrix–matrix product is attempted. In operator form the same step needs four solves and four sparse matrix–vector products, about 1 ms, and the sparsity of J_k is preserved throughout. For sparse problems the operator form is therefore not an optimisation but a precondition for using these schemes at all.

One caveat belongs here. The counts above assume that $\mathcal{F}'$ is available in closed form. If it is not, each Jacobian must be built from n directional differences or by algorithmic differentiation, at a cost of order n evaluations of $\mathcal{F}$ apiece, and the third Jacobian of PC8 then costs what the first two cost rather than being a small addition. In that regime the balance shifts: a two-Jacobian scheme of order six may be preferable to PC8 of order eight whenever $\mathcal{F}$ is expensive to evaluate, and the operator form remains the one improvement that is free in all cases. The same caveat applies to $\mathcal{F}'(\tilde{z}_k)$ specifically, which is required only to accuracy $O(h^3)$ by (27), so a cruder approximation than for $\mathcal{F}'(x_k)$ suffices; quantifying how crude is a question we leave open.

Table 2 collects the counts. PC8 pays one extra Jacobian relative to the cheapest two-Jacobian scheme and nothing else: the same single factorization, and $O(n^2)$ arithmetic for everything that

follows. When $\mathcal{F}'$ is available in closed form and n is large, the factorization dominates, and a step of PC8 costs about what a step of Newton's method costs.

7. Numerical experiments

All computations were performed in Python. The convergence tests use `mpmath` with 2000 significant digits, so that residuals of size 10^{-300} carry full accuracy; the timings use `numpy` and `scipy` in double precision on a single core. The test systems are

$$\begin{aligned}
&\text{S1:}\quad x_i - \cos\!\Big(2x_i - \textstyle\sum_{j=1}^{4} x_j\Big) = 0, \qquad i = 1,\dots,n;\\
&\text{S2:}\quad x_{i+1}x_i^2 - 1 = 0 \ (i<n), \qquad x_1 x_n^2 - 1 = 0;\\
&\text{S3:}\quad u_{j-1} - 2u_j + u_{j+1} + \tfrac{a^2}{4}(u_{j+1}-u_{j-1})^2 + h^2 = 0, \quad j = 1,\dots,n,\ a = 2,
\end{aligned}$$

S3 arising from $u'' + a^2(u')^2 + 1 = 0$, $u(0) = u(1) = 0$ on a uniform grid with $h = 1/(n+1)$, together with two small systems used to keep the high-precision runs cheap,

$$\begin{aligned}
&\text{S4:}\quad x_1^2 + x_2^2 - 4 = 0, \qquad e^{x_1} + x_2 - 1 = 0;\\
&\text{S5:}\quad x_1x_2 + x_3^2 - 3 = 0, \quad x_1^3 - x_2 + x_3 - 1 = 0, \quad x_1 + x_2^2 x_3 - 3 = 0,
\end{aligned}$$

and the discrete Chandrasekhar H-equation of radiative transfer,

$$\text{S6:}\quad x_i - \Big[1 - \frac{c}{2n}\sum_{j=1}^{n}\frac{\mu_i x_j}{\mu_i + \mu_j}\Big]^{-1} = 0, \qquad \mu_i = \frac{i-\frac{1}{2}}{n}, \quad c = 0.9, \tag{33}$$

which we include because its Jacobian is genuinely dense, so that the cost model of Section 6 applies without qualification, and the Bratu problem $u'' + \lambda e^u = 0$, $u(0) = u(1) = 0$, discretised by central differences as

$$\text{S7:}\quad u_{j-1} - 2u_j + u_{j+1} + h^2\lambda e^{u_j} = 0, \qquad j = 1,\dots,n, \quad \lambda = 2, \tag{34}$$

a standard semilinear elliptic problem. Neither its solution nor its structure falls under Proposition 4.6 or Proposition 4.7, so the barrier applies to it with no caveat.

7.1. Measuring the order.

A computed order of convergence obtained from three or four iterates [5] is a blunt instrument: when the asymptotic constant is small it can return a value one or two units above the truth, and it was in fact the discrepancy between such estimates and the local behaviour that first drew our attention to the problem. We therefore measure the local order of the one-step map. For $x_t = x^* + t\,d$ with a fixed generic unit vector d, and Φ one step of the method under test, we take

$$p \approx \frac{\log\big(\|\mathcal{F}(\Phi(x_{t_2}))\|/\|\mathcal{F}(\Phi(x_{t_1}))\|\big)}{\log\big(\|\mathcal{F}(x_{t_2})\|/\|\mathcal{F}(x_{t_1})\|\big)}, \qquad t_1 = 10^{-3},\ t_2 = 10^{-4}. \tag{35}$$

By a *generic* direction we mean one lying outside the proper algebraic subset of the unit sphere on which the leading error coefficient degenerates; Lemma 4.3 and Proposition 4.7 describe the two ways that degeneracy arises in practice. Operationally we fix one d whose entries stand in

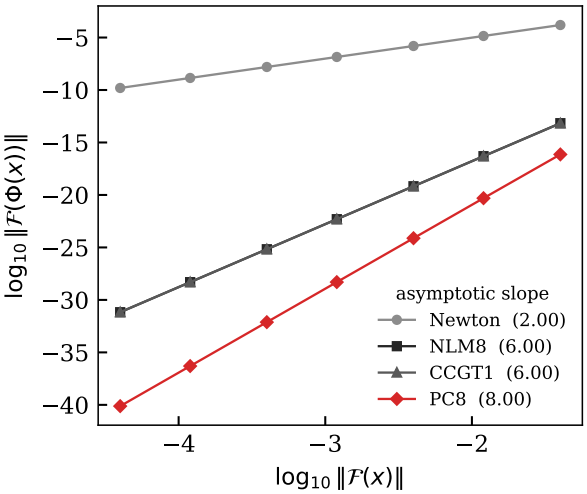


Figure 2. One step of each method applied to $x = x^* + t\,d$ on S5, with d a fixed generic direction and t decreasing from 10^{-2} to 10^{-5}. The slope of each curve is the order; the value in the legend is measured from the last pair of points. The two-Jacobian schemes lie on lines of slope six, PC8 on a line of slope eight.

no special relation to x^* and check that the measured slope is unchanged when d is replaced by other such choices, which it was in every case reported below.

On methods of known order this returns 2.0003 for Newton's method, 4.0010 for the two-step scheme with $\bar{\tau}_k = 2I - t_k$, and 8.0013 for three composed Newton steps, so the estimator is reliable to about two decimal places at these values.

Table 3 applies (35) to Newton's method (NM), to five weights from the literature, and to PC8. Four of them are labelled NLM8, CCGT1, CCGT2 and ZMO2 following [14, Table 1], where they are collected and attributed to [3, 10]; the fifth, NM8, is the scheme of [13] discussed in Proposition 4.8, obtained independently of that family.

Table 3. Local order (35) along a generic direction.

system	NM	NLM8	CCGT1	CCGT2	ZMO2	NM8	PC8
S4 (n=2)	2.000	6.000	6.001	6.001	6.001	6.000	**8.002**
S5 (n=3)	2.000	6.000	5.999	5.999	5.999	6.000	**7.997**
S1 (n=12)	1.999	5.984	5.997	5.995	5.993	5.984	**7.993**
S3 (n=12)	2.000	6.000	6.001	6.001	6.000	6.000	**8.003**
S6 (n=12)	2.000	5.973	5.998	5.995	5.991	5.973	**8.001**
S7 (n=12)	2.000	6.000	6.000	6.000	6.000	6.000	**8.001**

Figure 2 shows the same measurement graphically. The agreement with Theorem 4.4 is close: six for every weight that uses two Jacobians, eight for PC8, on systems of dimension two to

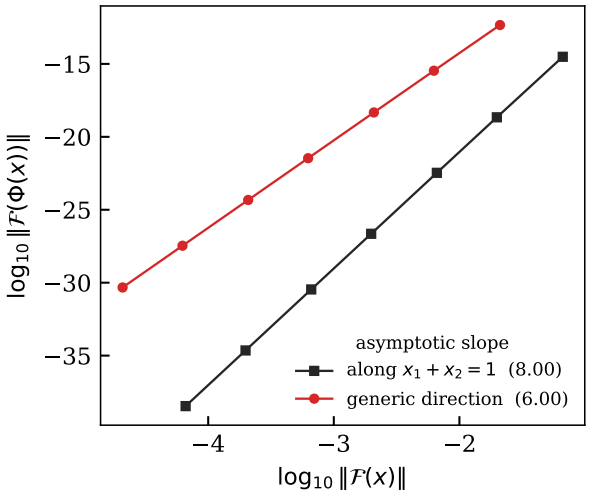


Figure 3. The same method, NLM8, on Example 1 of [10], approached from two directions. Along the line $x_1 + x_2 = 1$ fixed by the affine second equation the slope is eight; along a generic direction it is six. No property of the method changes between the two curves, only the set the iterate is allowed to move in.

twelve and of quite different character. The five weights come from four groups of authors who arrived at them by four different constructions — weight functions, a generating-function method, sufficient conditions on parameter matrices, and undetermined parameters — which is what one expects if the limitation belongs to the shape of (1) rather than to any particular derivation.

7.2. **Starting vectors that hide the barrier.** Proposition 4.6 predicts that a constant starting vector on S1 produces a scalar iteration. Table 4 confirms it. The perturbation $x_0 = 0.75\,e + 0.06\,(-1)^i$ changes the distance to the solution by a few percent and the measured order by two units.

Table 4. Local order (35) on S1 ($n = 12$) for two starting vectors.

starting vector	**NLM8**	**CCGT1**	**CCGT2**	**ZMO2**	**PC8**
$x_0 = 0.75\,e$, constant	8.12	8.13	8.13	8.13	8.12
x_0 non-constant	6.13	6.15	6.14	6.06	**8.07**

Figure 3 shows the two approaches side by side; the only difference between the curves is the direction of approach, and it moves the slope by two. Constant initial vectors are the natural choice for a system built from one repeated formula, and for such systems they reduce an n-dimensional test to a one-dimensional one. Anyone measuring an order numerically on this kind of problem should perturb the starting vector first.

7.3. **Residuals.** Table 5 lists $\|\mathcal{F}(x_k)\|$ for the first three iterations from non-degenerate starting vectors.

Table 5. Residuals $\|\mathcal{F}(x_k)\|$, $k = 1, 2, 3$.

system	**method**	$k = 1$	$k = 2$	$k = 3$
S4	NLM8	$1.08 \cdot 10^{-5}$	$1.33 \cdot 10^{-40}$	$1.42 \cdot 10^{-250}$
	CCGT1	$1.01 \cdot 10^{-6}$	$3.02 \cdot 10^{-47}$	$1.12 \cdot 10^{-289}$
	ZMO2	$3.64 \cdot 10^{-6}$	$5.67 \cdot 10^{-44}$	$2.58 \cdot 10^{-269}$
	PC8	$5.74 \cdot 10^{-6}$	$2.78 \cdot 10^{-46}$	$\mathbf{1.04 \cdot 10^{-368}}$
S5	NLM8	$5.74 \cdot 10^{-6}$	$1.23 \cdot 10^{-33}$	$1.40 \cdot 10^{-201}$
	CCGT1	$5.56 \cdot 10^{-6}$	$5.23 \cdot 10^{-35}$	$1.77 \cdot 10^{-209}$
	ZMO2	$4.22 \cdot 10^{-6}$	$7.63 \cdot 10^{-35}$	$1.44 \cdot 10^{-208}$
	PC8	$3.98 \cdot 10^{-6}$	$8.32 \cdot 10^{-44}$	$\mathbf{2.51 \cdot 10^{-345}}$
S1	NLM8	$3.82 \cdot 10^{-6}$	$3.64 \cdot 10^{-36}$	$8.47 \cdot 10^{-218}$
	CCGT1	$3.98 \cdot 10^{-6}$	$1.84 \cdot 10^{-35}$	$2.60 \cdot 10^{-214}$
	ZMO2	$3.36 \cdot 10^{-7}$	$1.91 \cdot 10^{-44}$	$1.87 \cdot 10^{-265}$
	PC8	$6.02 \cdot 10^{-6}$	$1.06 \cdot 10^{-43}$	$\mathbf{2.23 \cdot 10^{-347}}$
S3	NLM8	$1.14 \cdot 10^{-3}$	$1.05 \cdot 10^{-21}$	$5.41 \cdot 10^{-166}$
	CCGT1	$5.71 \cdot 10^{-5}$	$3.30 \cdot 10^{-36}$	$\mathbf{2.22 \cdot 10^{-259}}$
	ZMO2	$4.15 \cdot 10^{-4}$	$3.71 \cdot 10^{-26}$	$1.54 \cdot 10^{-202}$
	PC8	$9.44 \cdot 10^{-4}$	$1.57 \cdot 10^{-22}$	$9.64 \cdot 10^{-173}$

Order is an asymptotic statement, and the last block of Table 5 shows what that qualification is worth in practice. On S3 the error constants of PC8 are unfavourable, and after three steps from a moderate starting vector CCGT1 has the smaller residual even though its local order on that system is 6.001 against 8.003; the crossover occurs later. We report this rather than select the problems on which the asymptotic advantage is already visible.

7.4. **Timings.** Table 6 gives the wall-clock cost of one iteration on a dense system $\mathcal{F}(x) = Ax + \frac{1}{10}\sin x - b$ with A random and diagonally dominant.

At $n = 2000$ a step of PC8, of order eight, costs 0.33 s, against 1.14 s for a scheme of order six implemented with its weight matrices formed. Most of the gain comes from the operator form and is available to any scheme of the family; the remaining difference between the second and fourth columns, about twenty per cent, is the price of the third Jacobian and of the extra solves, and buys two orders of convergence.

7.5. **Efficiency index, and when the extra order actually pays.** The efficiency index $\mathrm{CEI} = p^{1/C}$, with C the cost of Table 2, is uninformative here: because C grows like n^3, every

Table 6. Seconds per iteration, double precision, single core.

n	**formed**	**operator form**	**PC8**
200	0.0045	0.0013	0.0028
500	0.0420	0.0097	0.0116
1000	0.2100	0.0501	0.0641
2000	1.1379	0.2716	0.3258

method returns $1 + O(10^{-6})$ already at $n = 100$. We therefore report $\log p/C$, which orders the methods in the same way and remains legible. At $n = 100$ the values, in units of 10^{-6}, are 0.99 for Newton's method, 0.38 for a two-Jacobian scheme with its matrices formed, 2.14 for the same scheme in operator form and 2.20 for PC8. PC8 is the most efficient, but its margin over the operator form of an order-six scheme is only about three per cent: the third Jacobian and the extra solves nearly consume what the two extra orders return.

The honest reading of that number is confirmed by a time-to-tolerance test. Table 7 solves S6 to $\|\mathcal{F}(x)\| < 10^{-13}$ in double precision.

Table 7. Iterations and time to reach $\|\mathcal{F}(x)\| < 10^{-13}$ on S6, double precision. NM denotes Newton's method.

	$n = 500$		$n = 1000$		$n = 2000$	
method	it.	s	it.	s	it.	s
Newton	5	0.042	5	0.139	5	0.789
order six, operator	2	0.025	2	0.086	2	0.456
PC8	2	0.028	2	0.105	2	0.510

Both high-order schemes reach machine accuracy in two iterations, so PC8 is about ten per cent slower than the order-six scheme it corrects. This is not an artefact of the example. In double precision a method of order six started reasonably close to the solution passes from 10^{-2} to below 10^{-13} in two steps, and no order above six can show its advantage before the arithmetic runs out. The two extra orders are worth having where the tolerance is genuinely tight — extended-precision computation, continuation methods in which the solution of one problem seeds the next, or ill-conditioned systems in which the residual stagnates — and the correctness of the order claim matters in all cases, since a method advertised as eighth order is often selected on that basis. What is unconditionally worth having at $n \gtrsim 500$ is the operator form, which costs nothing and saves a factor of three to five.

8. Concluding remarks

What decides the order of a frozen-Jacobian three-step method is how much of $\mathcal{F}''$ it gets to see. Two Jacobians, at x_k and at the Newton predictor, see it only along u_k; the middle and outer steps both need it along $Q_k u_k$ as well, and the shortfall costs exactly two orders. The fix asks only for that missing direction: evaluate the third Jacobian at the fourth-order predictor, where it is exposed, and keep every weight in operator form so that the extra information costs $O(n^2)$ rather than $O(n^3)$.

We have deliberately kept to local convergence under the smoothness hypotheses of Section 2. A convergence-ball analysis, bounding the radius in terms of Lipschitz constants for $\mathcal{F}'$ and $\mathcal{F}''$ in the manner now standard for Newton-type schemes in Banach space, would say how much the third Jacobian costs in domain of attraction, and we expect the answer to be unfavourable to PC8; establishing it properly is a separate undertaking and we have not attempted it here.

Three further questions seem worth pursuing. The first concerns divided differences. Amiri, Cordero, Darvishi and Torregrosa [1] replace the Jacobians of a scheme by the divided difference $[x + G(x), x; \mathcal{F}]$ with $G(x) = (f_1(x)^m, \ldots, f_n(x)^m)$, and prove that this preserves the error equation of the original method exactly, so a Jacobian-free NLM8 inherits whatever order NLM8 itself has. On a generic system that is order six, not eight, and our own runs of their scheme confirm it: the local order is 5.998 for $m = 2$ and 6.000 for $m = 3, 4$, matching the frozen-Jacobian NLM8 to three figures. Their construction is thus orthogonal to the present one: it removes the Jacobian at fixed order, whereas we add a directional derivative to raise the order, and combining the two — recovering R_k itself from a divided difference rather than a third Jacobian, which would cut the evaluation count from $3n^2$ to about $2n^2$ — is what we have not yet managed. The plain componentwise difference does not carry the second-order directional information R_k needs in a usable form; a symmetrised or higher-m variant may, at a cost we have not weighed. The second question is whether the barrier extends to schemes with more steps: the mechanism, that each new Jacobian exposes exactly one new direction, suggests a general statement relating the number of Jacobians to the attainable order. The third is the error constants, which Table 5 shows to be far from optimal for PC8; the conditions of Theorem 5.1 determine $\bar{\tau}_k$ and α_k only modulo $O(h^3)$, and the freedom left over could be used to reduce them.

Acknowledgements: The authors thank the referees for their careful reading and for several suggestions that improved the presentation. This work is supported by Buriram Rajabhat University. The scripts that generate every table and figure in this paper are provided as supplementary material.

Competing interests: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] A.R. Amiri, A. Cordero, M.T. Darvishi, J.R. Torregrosa, Preserving the order of convergence: Low-complexity Jacobian-free iterative schemes for solving nonlinear systems, J. Comput. Appl. Math. 337 (2018), 87–97. https://doi.org/10.1016/j.cam.2018.01.004.

[2] R. Behl, H. Arora, CMMSE: A novel scheme having seventh-order convergence for nonlinear systems, J. Comput. Appl. Math. 404 (2022), 113301. https://doi.org/10.1016/j.cam.2020.113301.

[3] A. Cordero, E. Gómez, J.R. Torregrosa, Efficient high-order iterative methods for solving nonlinear systems and their application on heat conduction problems, Complexity 2017 (2017), 6457532. https://doi.org/10.1155/2017/6457532.

[4] S. Huang, A. Rafiq, M.R. Shahzad, F. Ali, New higher order iterative methods for solving nonlinear equations, Hacet. J. Math. Stat. 47 (2018), 77–91. https://doi.org/10.15672/HJMS.2017.449.

[5] L.O. Jay, A note on Q-order of convergence, BIT Numer. Math. 41 (2001), 422–429. https://doi.org/10.1023/A:1021902825707.

[6] H.T. Kung, J.F. Traub, Optimal order of one-point and multipoint iteration, J. ACM 21 (1974), 643–651. https://doi.org/10.1145/321850.321860.

[7] A. Nadeem, F. Ali, J.-H. He, New optimal fourth-order iterative method based on linear combination technique, Hacet. J. Math. Stat. 50 (2021), 1692–1708. https://doi.org/10.15672/hujms.909721.

[8] J.M. Ortega, W.C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.

[9] M.S. Petković, B. Neta, L.D. Petković, J. Džunić, Multipoint methods for solving nonlinear equations, Elsevier, Amsterdam, 2013.

[10] J.R. Sharma, H. Arora, Improved Newton-like methods for solving systems of nonlinear equations, SeMA J. 74 (2017), 147–163. https://doi.org/10.1007/s40324-016-0085-x.

[11] J.F. Traub, Iterative methods for the solution of equations, Prentice-Hall, Englewood Cliffs, 1964.

[12] M. Turkyilmazoglu, A simple algorithm for high order Newton iteration formulae and some new variants, Hacet. J. Math. Stat. 49 (2020), 425–438. https://doi.org/10.15672/hujms.459810.

[13] X. Wang, Fixed-point iterative method with eighth-order constructed by undetermined parameter technique for solving nonlinear systems, Symmetry 13 (2021), 863. https://doi.org/10.3390/sym13050863.

[14] T. Zhanlav, R. Mijiddorj, O. Khuder, A family of Newton-type methods with seventh and eighth-order of convergence for solving systems of nonlinear equations, Hacet. J. Math. Stat. 52 (2023), 1006–1021. https://doi.org/10.15672/hujms.1061471.