The effects of rounding errors in the nodes on barycentric interpolation

Walter F. Mascarenhas and André Pierro de Camargo

Abstract We analyze the effects of rounding errors in the nodes on polynomial barycentric interpolation. These errors are particularly relevant for the first barycentric formula with the Chebyshev points of the second kind. Here, we propose a method for reducing them.

Mathematics Subject Classification (2010) 65D05, 65G50

1 Introduction

Given nodes \( x_0 < x_1 < \cdots < x_n \), weights \( w_0, \ldots, w_n \), and an interval \([x^-, x^+]\), two formulae for barycentric interpolation of a function \( f : [x^-, x^+] \to \mathbb{R} \) are considered. The first one is

\[
p(x; x, y, w) := \left( \prod_{k=0}^{n} (x - x_k) \right) \sum_{k=0}^{n} \frac{w_k y_k}{x - x_k},
\]

where \( y_k = f(x_k) \). The second one is

\[
q(x; x, y, w) := \sum_{k=0}^{n} \frac{w_k y_k}{x - x_k} / \left( \sum_{k=0}^{n} \frac{w_k}{x - x_k} \right).
\]

The first formula is a polynomial for any weights, but it only interpolates \( f \) if

\[
w_k = \lambda_k(x) := \frac{1}{\prod_{j \neq k} (x_k - x_j)}.
\]

The second formula always interpolates \( f \) at the nodes \( x_k \), but it is a polynomial for all \( y \) only if the weights are of the form \( w_k = \kappa_n \lambda_k(x) \), where \( \kappa_n \) is independent of \( k \).
Historically, Taylor \cite{17} and Salzer \cite{16} considered weights for which the second formula is a polynomial. The recent literature is also concerned with strictly rational second formulae \cite{1,2,3,6}, and we address this case in \cite{13}. Here, we focus on polynomial interpolation, specially in the classical case considered by Salzer, in which we interpolate at the Chebyshev points of the second kind.

In theory, barycentric interpolation at the Chebyshev points leads to accurate results. In practice, the first barycentric formula with these nodes suffers from accuracy problems when the number of nodes is large. In order to understand these problems we must consider the steps outlined in Figure 1:

Abstract function $f(x)$

**Step I: Abstract interpolation**

- Error I: Approximation theory

Abstract interpolant $a(x, \hat{x}, y, \hat{w})$

- $x =$ exact nodes
- $w =$ exact weights

**Step II: Finite precision representation of $b$.**

- Error II: Given by how $x, f(x)$ and $w$ are rounded.

In practice we use $a(x, \hat{x}, y, \hat{w})$

- $\hat{x} =$ rounded nodes
- $y =$ rounded $f(x)$
- $\hat{w} =$ rounded weights

**Step III: Evaluation of $a(x, \hat{x}, y, \hat{w})$**

- Error III: Usual stability analysis

Final result $f(x) \approx f_l(a(x, \hat{x}, y, \hat{w}))$

The overall error is a combination of the errors in the three steps.

Fig. 1: The overall error in interpolation. In this article $a$ is either $p$ in (1) or $q$ in (2).

The errors in Step II are not considered to their full extent in the current literature. For instance, \cite{9} takes into account the rounding errors in the evaluation of (3), but it does not consider weights obtained from closed form expressions evaluated at exact nodes, as in Salzer’s case. Table \cite{1} compares the errors introduced by Step II and by Step III for usual implementations of formulae (1) and (2) in Salzer’s case over two sets of points $X_{-1,n}$ and $X_{0,n}$ described in appendix \cite{3}. It shows that the errors introduced by Step II can be larger than those introduced by Step III for both formulae. The entries in bold face in Table \cite{1} highlight Step II errors that are much larger than Step III errors for the first formula.

| $n+1$ | First Formula | Second Formula |
|-------|---------------|---------------|
| $f(x) = \cos x$ | $f(x) = \cos(10^4 x)$ | $f(x) = \cos x$ | $f(x) = \cos(10^4 x)$ |
| $x \in X_{-1,n}$ | $x \in X_{-1,n}$ | $x \in X_{0,n}$ | $x \in X_{0,n}$ |
| $10^3$ | 1.4e+2 | 7.8e-1 | 4.4e-2 | 3.7e-2 |
| $10^4$ | 4.4e+3 | 1.0e-1 | critical | critical |
| $10^5$ | 1.5e+5 | 5.2e-1 | 5.6e-3 | 3.2e-3 | 6.6 | 8.9e-2 |
| $10^6$ | 8.4e+6 | 7.1e-2 | 6.6e-2 | 0.9e-3 | 1.0e-3 | 4.0 | 1.2e-2 |

This article estimates the errors in Steps II and III for the two barycentric formulae and proposes an strategy to reduce these errors in practice. In the next section we describe an experiment with the sine function which corroborates the data in Table \cite{1} and present an overview of our results. In Section \cite{3} we present our notation,
and estimates for the order of magnitude of the parameters relevant to our analysis of Salzer’s case. Section 4 analyzes how rounding errors in the nodes affect polynomial interpolation. Section 5 estimates the backward and forward errors for the first formula, and Section 6 presents bounds on the forward errors for the second formula (here, we do not present bounds on the backward error for the second formula because it is discussed in detail in [13].) Our strategy for reducing the errors in Step II is presented in the last section. Appendix A proves the lemmas and theorems stated in the previous sections whereas Appendix B describes the numerical experiments on which our tables and figures are based.

2 Overview and motivation

The complete analysis of the stability of the barycentric formulae requires much attention to detail, and people guided by concrete examples will have a better chance of understanding the subtle points. For this reason, throughout the article we illustrate the use of our general results in the following specific situation:

**Salzer’s Case.** We consider floating point nodes \( \hat{x}^{c} \) obtained by rounding the abstract Chebyshev points of the second kind \( x^{c} \):

\[
x^{(c)}_{k} := -\cos\left(\frac{k\pi}{n}\right) \quad \text{and} \quad \hat{x}^{(c)}_{k} := \text{rounded}\left(x^{(c)}_{k}\right) := \text{fl}\left(x^{(c)}_{k}\right).
\]

The weights used in computation are given in closed form by [16]. These weights are equivalent to \( \hat{w} = \lambda(x^{c}) \), for \( \lambda \) in [3], and we call them Salzer’s weights. We make conservative assumptions about \( n \) and the magnitude of the rounding errors. Formally, we suppose that the nodes are rounded as usual and

\[
10 \leq n \leq 2,000,000 \quad \text{and} \quad \|\hat{x}^{c} - x^{c}\|_{\infty} \leq 4.6 \times 10^{-16}.
\]

Salzer’s case is relevant first because of its practical importance, second because it shows clearly that the concern with the perturbation in the nodes is not futile. In fact, if we neglect these errors then we can underestimate the errors in the first formula by orders of magnitude in this case. Therefore, in order to fully understand the accuracy of the barycentric formulae \( p \) and \( q \) in (1) and (2), one must be aware of the differences between their variations (a), (b) and (c) below, in which \( \hat{w} \) are the weights used in computation (for instance, in Salzer’s case \( \hat{w} = \lambda(x^{c}) \) and [9] considers \( \hat{w} = \text{fl}(\lambda(\hat{x})) \)).

\[
\begin{align*}
(a) \quad p(x; x, y, \lambda(x)) &= p(x; \hat{x}, y, \hat{w}) \\
(b) \quad q(x; x, y, \lambda(x)) &\neq q(x; \hat{x}, y, \hat{w}) \\
(c) \quad P(x; x, y) &\neq P(x; \hat{x}, y)
\end{align*}
\]

where \( P(x; x, y) := \) the \( n \)th degree polynomial that interpolates \( y \) at \( x \).
In variation (a) we consider the theoretical nodes, like the Chebyshev points of the four kinds. Usually, these theoretical nodes cannot be represented exactly in finite precision arithmetic and in practice we use rounded nodes instead, and these rounded nodes are considered in variations (b) and (c). The rounded nodes do not have all the theoretical properties of the exact ones, like the orthogonality of the corresponding polynomials with respect to convenient inner products or neat closed form expressions relating them. Unfortunately, the advantages given by these theoretical properties may be illusory for large $n$, and we may be subject to subtle side effects when we apply results derived for the exact nodes to the rounded ones. For instance, if we use Fourier transforms based on the exact nodes to obtain weights for the barycentric formulae then we obtain weights like Salzer’s. However, in the following experiment with the first formula applied to the sine function the maximum forward error

$$\max_{\text{trial points } x} | \sin(x) - p(x; \tilde{x}, \sin(\tilde{x}), \lambda(\tilde{x})) |$$

(7)

corresponding to the Salzer’s weights $\lambda(\tilde{x})$ was about 650 times larger than the error corresponding to the numerical weights $\text{fl}(\lambda(\tilde{x}))$ for $n = 1000$.

The set of trial points in (7), and the details of the experiment in Figure 2, are presented in appendix B; the nodes $\tilde{x}$ in this experiment were computed with machine precision $\epsilon \approx 2.3 \times 10^{-16}$, and the straight lines in this plot were obtained by the least squares method. The straight line for Salzer’s weights shows that, in this particular experiment, the corresponding errors grow like $0.1\epsilon n^2$. On the other hand, the errors incurred when using rounded nodes in combination with the Numerical weights are in better agreement with the $O(\epsilon n)$ upper bounds presented in [9]. In Section 5 we explain this $O(\epsilon n^2) \times O(\epsilon n)$ discrepancy in the order of magnitude of the forward errors for the first formula. In summary, we show that, for large $n$, the maximum forward error for the first formula in Steps II and III in Salzer’s case is well described by

![Fig. 2: log$_{10}$ of the maximum error in the approximation of $f(x) = \sin x$ in $[-1, 1]$ by the first barycentric formula with rounded nodes $\tilde{x}$ and (i) Salzer’s weights, $\lambda(\tilde{x})$, and (ii) the weights corresponding to the rounded nodes, $\text{fl}(\lambda(\tilde{x}))$, which we call by Numerical weights.](image-url)
where $yz \in \text{terms of the machine precision}$

In the following sections we study them in detail, and explain how they can be bounded in the first and the second barycentric formula in Salzer’s case. For this reason, in the key factors for the understanding of the maximum forward errors in Step II for the rounded nodes in combination with the Numerical weights is of order $\varepsilon n$ (see Tables 1 and 2 in [13], in which $\zeta_k = -z_k/(1+z_k) \approx -z_k$).

For the second formula, we show that the forward error in Step II can be estimated via the Error Polynomial $E(x; \hat{x}, y, z)$, which is given by

$$E(x; \hat{x}, y, z) := P(x; \hat{x}, y, z) - P(x; y, z)P(x; \hat{x}, z),$$

where $yz$ is the vector with entries $(yz)_k = y_kz_k$. The $z_k$ and the Error Polynomial are the key factors for the understanding of the maximum forward errors in Step II for the first and the second barycentric formula in Salzer’s case. For this reason, in the following sections we study them in detail, and explain how they can be bounded in terms of the machine precision $\varepsilon$.

$L :=$ Lipschitz constant of the function $f$ we are interpolating,

the Lebesgue constant

$$\Lambda := \Lambda_{x, x^+} := \sup_{x \in [x, x^+]} \frac{|P(x; x, y)|}{\|y\|_\infty},$$

and terms related to the node spacing. Our conclusions are summarized by the following diagram, in which $P^*$ is the best polynomial approximation of $f$:

$$\begin{array}{c}
p(x; x, y, \lambda(x)) = P(x; x, y) = q(x; x, y, \lambda(x)) \\
p(x; \hat{x}, \lambda(\hat{x})) \\
Λ (L\|x - x\|_\infty + \|f - P^*\|_\infty) \\
Λ \|yz\|_\infty \\
E(x; \hat{x}, y, z) \\
p(x; \hat{x}, y, \lambda(\hat{x})) = P(x; \hat{x}, y) = q(x; \hat{x}, y, \lambda(\hat{x}))
\end{array}$$

Fig. 3: The forward errors for the formulae $p$ and $q$ in Step II. The values in the edges of this diagram are upper bounds on the order of magnitude of the maximum difference between their end points. These bounds hold under technical conditions described in the next sections.

In the following sections we look carefully at the differences

$$p(x; \hat{x}, y, \lambda(\hat{x})) - p(x; \hat{x}, y, \hat{\lambda}(\hat{x})) \quad \text{and} \quad q(x; \hat{x}, y, \lambda(\hat{x})) - q(x; \hat{x}, y, \hat{\lambda}(\hat{x}))$$

shown in the lower edges in Figure [3] and justify the estimates presented in these edges. Using the vertical edge in this figure and the triangle inequality, we can bound

$$p(x; \hat{x}, y, \lambda(\hat{x})) - p(x; x, y, \lambda(\hat{x})) \quad \text{and} \quad q(x; \hat{x}, y, \lambda(\hat{x})) - q(x; x, y, \lambda(\hat{x})), $$
and in Section 4 we present bounds on the difference $P(x; y) - P(x; \hat{x}, y)$ corresponding to this vertical edge.

We end this overview emphasizing that the Lipschitz constant plays an important role in the accuracy of Step II for the second formula. In Salzer's case, the function

$$R(x; \hat{x}) := \frac{\gamma^{n-1}}{\sqrt{1-x^2}} \prod_{k=0}^{n} (x - \hat{x}_k)$$

is highly oscillating, with maximum absolute value close to 1, because $R(x; x^c) = -\sin(n \arccos x)$. The function $R(x; \hat{x})$ has simple zeros in $(-1, 1)$, and all of them are zeros of $E(x; \hat{x}, y, z)$. It follows that

$$Q(x; \hat{x}, y, z) := \frac{E(x; \hat{x}, y, z)}{R(x; \hat{x})}$$

is also a smooth function. This leads to the decomposition $E(x; \hat{x}, y, z) = R(x; \hat{x}) \times Q(x; \hat{x}, y, z)$. The factor $Q(x; \hat{x}, y, z)$ depends on the function $f$ which we are interpolating. Figure 4 illustrates graphically the decomposition $E = R \times Q$ for $f(x) = \cos(10x)$ in Salzer’s case.

![Fig. 4: The factor $Q$ in (9) in Salzer’s case with $n + 1 = 100$ nodes and $f(x) = \cos(10x)$. Notice that $Q$ is $O(\varepsilon)$ and does not oscillate much.](image)

Figure 5 shows that as we increase the Lipschitz constant we increase the amplitude as well as the frequency of the errors. In the extreme case given by the Lagrange polynomials with nodes $\hat{x}$, the Lipschitz constant is of order $n^2$, and the bounds presented in this article are not encouraging. In fact, [13] shows that the maximum backward error for the second formula is of order $\varepsilon n^2$ for Lagrange polynomials in Salzer’s case.

3 Notation and conventions

Throughout the article, we consider intervals $[\hat{x}^-, \hat{x}^+]$, $[x^-, x^+]$ nodes $\hat{x}$ and $x$ and weights $\hat{w}$ and $w$, and the readers may find it convenient to have a copy of the next
We measure the difference between \( x \) and \( \hat{x} \) in terms of

\[
x_k < x_{k+1}, \quad \hat{w}_k \neq 0 \quad \text{and} \quad w_k \neq 0, \quad (10)
\]

\[
x_k \in (x^-, x^+) \quad \text{if and only if} \quad \hat{x}_k \in (\hat{x}^-, \hat{x}^+),
\]

\[
x_k = x^- \quad \text{if and only if} \quad \hat{x}_k = \hat{x}^-,
\]

\[
x_k = x^+ \quad \text{if and only if} \quad \hat{x}_k = \hat{x}^+,
\]

\[
k^- \text{ is the smallest } k \text{ such that } x_k > x^-,
\]

\[
k^+ \text{ is the largest } k \text{ such that } x_k < x^+ \quad \text{and} \quad k^- \leq k^+. \quad (13)
\]

We measure the difference between \( x \) and \( \hat{x} \) in terms of

\[
\delta_{kj} := \delta_{kj}(\hat{x}, x) := 0 \quad \text{and} \quad \delta_{jk} := \delta_{jk}(\hat{x}, x) := \frac{x_j - x_k}{\hat{x}_j - \hat{x}_k} - 1, \quad (14)
\]

\[
\delta^-_j := \delta^-_{jk}(\hat{x}^-, \hat{x}, x^-) := \frac{x^- - x_j}{\hat{x}^- - \hat{x}_j} - 1, \quad \delta^+_j := \delta^+_{jk}(\hat{x}^+, \hat{x}, x^+) := \frac{x^+ - x_j}{\hat{x}^+ - \hat{x}_j} - 1,
\]

with the convention that \( \delta^-_j := 0 \) and \( \delta^+_j := 0 \) in the exceptional cases \( \hat{x}_j = \hat{x}^- \) and \( \hat{x}_j = \hat{x}^+ \). We combine the \( \delta^-_j \), \( \delta^+_j \), and \( \delta^+_j \) in

\[
\delta := \max_{0 \leq j, k \leq n} \left\{ |\delta^-_j|, |\delta_k|, |\delta^+_j| \right\}. \quad (15)
\]

We also measure the errors in the nodes by

\[
\Delta^- := \sum_{j=0}^{n} \max \left\{ |\delta^-_j|, |\delta^-_{jk}| \right\}, \quad \Delta_k := \sum_{j<k} \max \left\{ |\delta^+_j|, |\delta^+_{jk}| \right\},
\]

\[
\Delta^+ := \sum_{j=0}^{n} \max \left\{ |\delta^+_{jk}|, |\delta^+_{jk+1}| \right\}, \quad \Delta := \max_{k^- \leq k \leq \min\{k^+, n-1\}} \{ \Delta^-, \Delta_k, \Delta^+ \}. \quad (16)
\]

(\( \Delta_k \) is defined for \( 0 \leq k < n \).) The differences between the weights \( \hat{w} \) used in computation and the weights \( \lambda_k(\hat{x}) \) corresponding to the rounded nodes according to \( \hat{x} \) are measured by

\[
z_k := z_k(\hat{x}, \hat{w}) := \frac{\hat{w}_k - \lambda_k(\hat{x})}{\lambda_k(\hat{x})} \quad \text{and} \quad \z_k := z_k(w, \hat{w}) = \frac{w_k - \hat{w}_k}{w_k}, \quad (17)
\]
in which usually \( w_k = \lambda_k(\hat{x}) \). The Lebesgue constant is defined as

\[
\Lambda := \Lambda_{\hat{x} - \hat{x}^+} := \sup_{x \in [\hat{x}^-, \hat{x}^+]} \frac{|P(x; x, y)|}{\|y\|_{\infty}},
\]

(18)

where \( P(x; x, y) \) is the \( n \)th degree polynomial that interpolates \( y \) at \( x \). Note that this definition implies that if \( Q \) is a \( n \)th degree polynomial then

\[
|Q(x_k)| \leq M \quad \text{for} \quad 0 \leq k \leq n \quad \Rightarrow \quad |Q(x)| \leq \Lambda_{\hat{x} - \hat{x}^+} M \quad \text{for} \quad x \in [\hat{x}^-, \hat{x}^+],
\]

(19)

because \( Q(x_k) = P(x; x, Q(x)) \). The \( k \)th Lagrange polynomial with nodes \( x \) is given by

\[
\ell_k(x; x) := \lambda_k(x) \prod_{j \neq k} (x - x_j).
\]

(20)

The \( z_k, \zeta_k \) and \( \delta_{jk} \) are related by the following lemma:

**Lemma 1** If \( a_k := \sum_{j \neq k} |\delta_{jk}| < 1 \) then \( z_k = z_k(\hat{x}, \lambda_k(x)) \) and \( \zeta_k = \zeta_k(\lambda_k(\hat{x}), \lambda_k(x)) \) satisfy

\[
0 \leq z_k + \sum_{j \neq k} |\delta_{jk}| \leq \frac{a_k}{1 - a_k}, \quad |z_k| \leq \frac{a_k}{1 - a_k} \quad \text{and} \quad |\zeta_k| \leq \frac{a_k}{1 - a_k}.
\]

(21)

Using \( \Delta \), we can bound products of the form \( \prod (x - x_k) / (\hat{x} - \hat{x}_k) \):

**Lemma 2** Under the conditions (10)–(13), if \( \Delta < 1 \) then for every \( \hat{x} \in [\hat{x}^-, \hat{x}^+] \) there exists \( x(\hat{x}) \in [\hat{x}^-, \hat{x}^+] \) such that

\[
|x(\hat{x}) - \hat{x}| \leq \max \left\{ \|x - \hat{x}\|_{\infty}, \|x^- - \hat{x}^-\|, \|x^+ - \hat{x}^+\| \right\}
\]

(22)

and, for the same \( x(\hat{x}) \) and every \( K \subset \{0, 1, \ldots, n\} \), there exists \( \beta_K \) such that

\[
|\beta_K| \leq \frac{\Delta}{1 - \Delta} \quad \text{and} \quad \prod_{k \in K} (\hat{x} - \hat{x}_k) = (1 + \beta_K) \prod_{k \in K} (x(\hat{x}) - x_k).
\]

(23)

In Salzer’s case, we can bound \( \delta, \Delta, z \) and \( \zeta \) in terms of \( n \) and the rounding errors in the nodes, and show that our theory applies even to \( n \) in the million range:

**Lemma 3** In Salzer’s case we have the bounds in Table 2 where \( z = z(\hat{x}, \lambda(x)) \) and \( \zeta = \zeta(\lambda(\hat{x}), \lambda(x)) \):
The effects of rounding errors in the nodes on barycentric interpolation

Table 2: Upper bounds for Salzer’s case

|       | upper bound for n | absolute upper bound |
|-------|-------------------|----------------------|
| δ     | $0.40897 | \|x' - x\|_{\infty} n^2$ | $1.8813 \times 10^{-16} n^2$ | $7.5252 \times 10^{-4}$ |
| Δ     | $2.7267 | \|x' - x\|_{\infty} n^2$ | $1.2543 \times 10^{-15} n^2$ | $5.0172 \times 10^{-3}$ |
| $\max \sum [\|z\|_m]$ | $2.4502 | \|x' - x\|_{\infty} n^2$ | $1.1271 \times 10^{-15} n^2$ | $4.3084 \times 10^{-3}$ |
| $\|\zeta\|_m$ | $2.4624 | \|x' - x\|_{\infty} n^2$ | $1.1328 \times 10^{-15} n^2$ | $4.5312 \times 10^{-3}$ |
| $\|z\|_4$ | $0.71094 | \|x' - x\|_{\infty} (2.9 + \log n)^2$ | $3.2704 \times 10^{-16} (2.9 + \log n)^2$ | $3.9646 \times 10^{-1}$ |

4 Perturbations in the nodes of Polynomials

There are at least four reasonable concepts of “the interpolating polynomial of $f$” when we take into account the rounding errors in the nodes: $P(x; x, f(x)), P(x; x, \hat{f}(\hat{x})), P(x; \hat{x}, f(x))$ and $P(x; \hat{x}, \hat{f}(\hat{x}))$. The difference $\hat{f}(\hat{x}) - f(x)$ is of order $L \|x - \hat{x}\|_{\infty}$, where $L$ is $f$’s Lipschitz constant. This leads to a difference $P(x; x, \hat{f}(\hat{x})) - P(x; x, f(x))$ of order $\Lambda_{x-x} \|x - \hat{x}\|_{\infty}$, and this section shows that when $f$ is well approximated by polynomials this is the overall order of magnitude of the difference $P(x; x, y) - P(x; \hat{x}, y)$, as suggested in the vertical edge in Figure 3.

We show also that the usual error estimate for Lagrange interpolation, namely,

$$f^{(n+1)}(\xi) \frac{1}{(n+1)!} \prod_{k=0}^{n} (x - x_k),$$

(24)

does not change much when we replace $x$ by $\hat{x}$ and the $\Delta$ in (16) is small. Therefore, $P(x; \hat{x}, \hat{f}(\hat{x}))$ is as good an approximation of $f$ as $P(x; x, f(x))$ if we consider (24) as a measure of the degree of approximation. Moreover, we show that when $\Delta$ is small the Lebesgue constant with respect to the nodes $\hat{x}$ is roughly the same as the Lebesgue constant corresponding to the nodes $x$. The overall conclusion of this section is that it is reasonable to consider the differences

$$p(x; \hat{x}, y, \hat{w}) - p(x; \hat{x}, y, \hat{w}), \quad q(x; \hat{x}, y, \hat{w}) - q(x; \hat{x}, y, \hat{w})$$

as measures of the errors in Step II for the first and second formula, as in Salzer’s case. The informal statements above are formalized by the following results:

**Lemma 4** Under the conditions $[10] - [13]$, if $z = z(\hat{x}, \hat{x}(x))$ satisfies

$$\delta < \frac{1 - \|z\|_{\infty}}{\Lambda_{x-x} x},$$

(25)

then

$$\Lambda_{x-x} x \leq \frac{1 + \delta}{1 - \|z\|_{\infty} - (\delta + \|z\|_{\infty}) \Lambda_{x-x} x}.$$

(26)

In particular, in Salzer’s case,

$$\Lambda_{x-x} \leq 1.0629 \Lambda_{x-x} \leq 0.67667 \log n + 1.0236 \quad \text{and} \quad \Lambda_{x-x} \leq 10.841.$$  

(27)
Lemma 5 Under the conditions (10)–(13), if $\Delta < 1$ then the usual estimate \([24]\) for \([\hat{x}^-, \hat{x}^+]\) and \(\hat{x}\) is not much larger than the same estimate for \([x^-, x^+]\) and \(x\), namely,

$$
\max_{x \in [\hat{x}^-, \hat{x}^+] \setminus \{\hat{x}_k\}} \prod_{k=0}^n (\hat{x} - \hat{x}_k) \leq \frac{1}{1 - \Delta} \max_{x \in [x^-, x^+] \setminus \{x_k\}} \prod_{k=0}^n (x - x_k). \quad (28)
$$

In Salzer’s case,

$$
\max_{x \in [-1, 1]} \prod_{k=0}^n (\hat{x} - \hat{x}_k(c)) \leq \left(1 + 1.2602 \times 10^{-15} n^2\right) \max_{x \in [-1, 1]} \prod_{k=0}^n (x - x_k(c)) \quad (29)
$$

$$
\leq 2^{1-n} \left(1 + 1.2602 \times 10^{-15} n^2\right) \leq 2.0101 \times 2^{-n}.
$$

It is well known that the best possible value for the left hand side of (29), among all sets of nodes, is at most $2^{-n}$ (see \([16]\)). Therefore, the rounded nodes $\hat{x}$ are nearly optimal concerning the size of the product in the bound \([24]\). More generally, Lemma 5 shows that, for reasonably rounded nodes, the accuracy for interpolation of variations (a) and (c) in \([6]\) cannot be distinguished solely on basis of the traditional estimate \([24]\). Finally, we present a theorem formalising the bound presented on the vertical edge in Figure 3.

Theorem 1 Suppose that $\hat{x}^- \geq x^-$ and $\hat{x}^+ \leq x^+$ and consider a function $f : \mathbb{R} \to \mathbb{R}$ with Lipschitz constant $L$ and such that $y_k = f(x_k)$. If $Q$ is a polynomial of degree at most $n$ and $x \in [\hat{x}^-, \hat{x}^+]$ then

$$
|P(x; \hat{x}, y) - P(x; x, y)| \leq L\Lambda_{\hat{x}^-, \hat{x}^+} \|\hat{x} - x\|_\infty + M \left(\Lambda_{\hat{x}^-, \hat{x}^+} + \Lambda_{\hat{x}^-, \hat{x}^+}\right), \quad (30)
$$

where

$$
M := \max_{x \in \{x_0, x_1, \ldots, x_n\} \setminus \{\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_n\}} |Q(x) - f(x)|. \quad (31)
$$

In Salzer’s case, for all $x \in [-1, 1]$, we have

$$
|P(x; \hat{x}, y) - P(x; x, y)| \leq (0.7 \log n + 1) L \|\hat{x} - x\|_\infty + M \left(2 + 1.4 \log n\right). \quad (32)
$$

5 Bounds for the first formula

Here we discuss the forward and backward stability of the first formula in Steps II and III. The first issue we address is the appropriate concept of backward stability when we allow for rounded nodes. We then present upper bounds on the error in Steps II and III, and lower bounds for Step II, for the first formula. In particular, we show that, when errors in the nodes are taken into account, the overall error in steps II and III in Salzer’s case is of order $\hat{e}^n n^2$ for this formula.

We now explain that when the nodes are perturbed we cannot restrict ourselves to perturbation in the function values in order to prove backward stability. In fact, let us consider the Lagrange polynomials, which can be written in first barycentric form as

$$
\ell_k(x; x) = p\left(x, x, e^{y_k}, \lambda(x)\right),
$$
where \( y_k = 1 \) and \( e^k \in \mathbb{R}^{n+1} \) is the vector with \( e^k_j = 1 \) and \( e^k_j = 0 \) for \( j \neq k \). If we were to consider only relative perturbations in \( y \), then given rounded nodes \( \hat{y} \) we would need to find \( \hat{y}_k = (1 + \beta_k) y_k = 1 + \beta_k \) such that

\[
p(x, \hat{y}, e^k, \lambda(x)) = p(x, x, e^k (1 + \beta_k), \lambda(x)).
\]

This equation leads to

\[
\beta_k = \prod_{i \neq k} \frac{x - \hat{x}_k}{x - x_k} - 1,
\]

and given an arbitrarily small \( \varepsilon > 0 \), we could take \( k = 1, \hat{x}_j = x_j \) for \( k \neq 1, \hat{x}_1 = x_1 + \varepsilon \) and obtain

\[
\beta_1 = \frac{x - x_1 - \varepsilon}{x - x_1} - 1 = \frac{-\varepsilon}{x - x_1}.
\]

We could then make \( \beta_1 \) arbitrarily large by taking \( x \) close enough to \( x_1 \). Therefore, we cannot build a backward stability theory for the first formula with perturbed nodes relying only on perturbations of the function values. On the other hand, the next theorem shows that we can get meaningful results if we allow for perturbations in \( x \):

**Theorem 2** Under the conditions (70), if \( \Delta < 1 \) and the machine precision \( \varepsilon \) is such that \( (3n + 5) \varepsilon < 1 \) then for every \( \hat{x} \in [\hat{x}^-, \hat{x}^+] \) and \( y \in \mathbb{R}^{n+1} \) there exists \( x \in [x^-, x^+] \) and \( \beta, \nu \in \mathbb{R}^{n+1} \) such that

\[
\|\beta\|_\infty \leq \Delta / (1 - \Delta), \quad \|\nu\|_\infty \leq \frac{(3n + 5) \varepsilon}{1 - (3n + 5) \varepsilon}
\]

and

\[
|\hat{x} - x| \leq \max \{\|\hat{x} - x\|_\infty, |\hat{x}^- - x^-|, |\hat{x}^+ - x^+| \}, \quad (33)
\]

and the vector \( \hat{y} \) with

\[
\hat{y}_k := y_k (1 + \beta_k) (1 + \nu_k) \left( 1 + \frac{\hat{w}_k - w_k}{w_k} \right) \quad (34)
\]

satisfies

\[
\varepsilon_l(p(\hat{x}, \hat{y}, w)) = p(x, x, \hat{y}, w). \quad (35)
\]

In words, Theorem 2 shows that the first formula is backward stable in Steps II and III, in the broader sense which allows also for perturbations in \( x \).

Our analysis of the forward stability of the first formula in Steps II and III is more complete than the analysis of the backward stability, because it also yields a lower bound on the error, which is described by the first inequality in equation (37) in the following theorem:

**Theorem 3** If \( \hat{x} \in [\hat{x}^-, \hat{x}^+] \) and the machine precision \( \varepsilon \) is such that \( (3n + 5) \varepsilon < 1 \) then the first formula \( p \) in (1) satisfies

\[
|\varepsilon_l(p(\hat{x}, \hat{y}, w)) - P(\hat{x}, \hat{y}, w)| \leq A_{\hat{x}^- \hat{x}^+ \hat{x}^+} \max_{0 \leq j \leq n} \left| \frac{y_j + (3n + 5) \varepsilon}{1 - (3n + 5) \varepsilon} \right|, \quad (36)
\]
for \( z = z(\mathbf{x}, \mathbf{w}) \). Moreover, for every \( k \) such that \( \hat{z}_k \in [\hat{x}^-, \hat{x}^+] \), we have

\[
|y_kz_k| (1 - (3n + 5)\varepsilon) \leq \frac{(3n + 5)\varepsilon|y_k|}{1 - (3n + 5)\varepsilon}
\]

\[
\leq \lim_{\hat{x} \to \hat{z}_k, \hat{z} \in [\hat{x}^-, \hat{x}^+] |f(p(\hat{x}, \mathbf{x}, \mathbf{w})) - P(\hat{x}, \mathbf{x}, \mathbf{y})|} \leq \left| y_k \frac{z_k + (3n + 5)\varepsilon}{1 - (3n + 5)\varepsilon} \right|
\]

(37)

In Salzer’s case, Lemma[5] shows that the upper bound (36) is of order \( \varepsilon n^2 \log n |y_j| \).

Of course, in itself, this upper bound does not imply that the backward errors will be of order \( \varepsilon n^2 \log n \|y\|_\infty \) in this case. In fact, in the usual situations in which \( x \) is not very close to the nodes, the backward error will be much smaller than the right hand side of (36). On the other hand, Table 1 in [13] provides strong empirical evidence that \( \|z\|_\infty \) is at least 0.01\( n \pi \). Therefore, the first inequality in (37) suggests that whenever the \( y_k \) corresponding to \( |z_k| \approx \|z\|_\infty \) is not small, it is likely that we will incur in errors of magnitude \( \varepsilon n^2 \|y\|_\infty \) when we evaluate the first barycentric formula for \( \hat{x} \) very close to \( \hat{z}_k \). For instance, when \( n = 10^6 \) and \( \varepsilon = 2.3 \times 10^{-16} \) we have

\[
0.01\varepsilon n^2 = 2.3 \times 10^{-12} = 2.3 \times 6 - 10^{-6},\text{ and this value agrees remarkably well with the maximum error of } 2 \times 10^{-6},
\]

for the function \( \cos(100x) \) presented in Tables 4 and 5 in Subsection 7.2.

6 Bounds for the second formula

In this section we bound the errors in Step II and III in Figure 1 for the second barycentric formula (2) in terms of the Error Polynomial in (8). The Error Polynomial is similar to a function presented by Werner [18], which expresses its results in terms of divided differences. We, on the other hand, use the Error Polynomial in combination with our bounds on \( z \) in order to have a unified picture for both barycentric formulae and to obtain more explicit bounds. The next theorem relates the Error Polynomial, the Lebesgue constant \( \Lambda_{\hat{x}^- \hat{x}^+} \), and the forward error in Step II for the second formula:

**Theorem 4** Consider \( z = z(\mathbf{x}, \mathbf{w}) \). If \( \hat{x} \in [\hat{x}^-, \hat{x}^+] \) and \( \Lambda_{\hat{x}^- \hat{x}^+} \|z(\hat{x}, \mathbf{w})\|_\infty < 1 \) then the second formula \( q \) in (2) satisfies

\[
\frac{|E(\hat{x}; \mathbf{x}, \mathbf{y}, z)|}{1 + \Lambda_{\hat{x}^- \hat{x}^+} \|z\|_\infty} \leq |q(\hat{x}; \mathbf{x}, \mathbf{y}, z) - P(\hat{x}; \mathbf{x}, \mathbf{y})| \leq \frac{|E(\hat{x}; \mathbf{x}, \mathbf{y}, z)|}{1 - \Lambda_{\hat{x}^- \hat{x}^+} \|z\|_\infty}.
\]

(38)

We now present an empirical way to bound \( |E(\hat{x}; \mathbf{x}, \mathbf{y}, z)| \), and after this empirical bound we present a theoretical bound of order \( L|n\log^2 n \) on the forward error of the second barycentric formula. Given \( y \) and \( z \), we can compute bounds on the Error Polynomial in terms of its values at convenient points \( c_k \). In fact, given points \( \{c_0, c_1, \ldots, c_m\} \) we define

\[
C := \{\hat{z}_k, 0 \leq k \leq n\} \bigcup \{c_k, 1 \leq k \leq m\}
\]

(39)
and the identity $E(\times; \hat{\mathbf{x}}, \mathbf{y}, \mathbf{z}) = 0$ and the definition of Lebesgue constant \[18\] yield

$$\max_{x \in |x - \hat{x}|} |E(x; \hat{x}, \mathbf{y}, \mathbf{z})| \leq \Lambda_{k} - x^{+} C \max_{1 \leq k \leq m} |E(c_{k}; \hat{x}, \mathbf{y}, \mathbf{z})|. \quad (40)$$

The right-hand side of \[40\] overestimates the left-hand side by at most a factor of \(\Lambda_{k} - x^{+} C\), and when we choose appropriate sets \(C\) this factor is of order \(\log n\). For instance, in Salzer’s case, if we choose \(C\) as the rounded Chebyshev nodes corresponding to \(2n\) then Lemma \[8\] shows that \(\Lambda_{k} - x^{+} C \leq 0.7 \log (2n) + 1.1\) for \(10 \leq n \leq 1,000,000\).

Equation \[40\] leads to computable bounds for the forward errors for relevant classes of functions. For instance, we can bound the error in Step II for a function \(f\) with a Lipschitz constant \(L\) by evaluating \(z\) and maximizing the linear functions of \(y\) given by

$$h_{j}(y) := E(c_{k}; \hat{x}, y, z) = \sum_{j=0}^{n} a_{k,j} y, \quad (41)$$

where the \(a_{k,j}\) are defined using \(\ell_{j}(x; \hat{x})\), the \(j\)-th Lagrange polynomial with nodes \(\hat{x}\),

$$a_{k,j} := a_{k,j}(C, z) := (z_{j} - P(c_{k}; \hat{x}, z)) \ell_{j}(c_{k}; \hat{x}), \quad (42)$$

subject to linear constraints of the form

$$y_{j} - y_{j-1} \leq L(\hat{x}_{j} - \hat{x}_{j-1}) \quad \text{and} \quad y_{j-1} - y_{j} \leq L(\hat{x}_{j} - \hat{x}_{j-1}). \quad (43)$$

The linear programming problem with objective function \[41\] and constraints \[43\] can be solved in closed form. Its solution leads to the following lemma:

**Lemma 6** Consider \(a_{k,j}\) in \[42\]. If \(y\) satisfies \[43\] for \(1 \leq j \leq n\) then

$$|E(c_{k}; \hat{x}, y, z)| \leq L \sum_{i=1}^{n} (\hat{x}_{i} - \hat{x}_{i-1}) \left| \sum_{j=i}^{n} a_{k,j} \right|. \quad (44)$$

This lemma and equation \[40\] lead to

$$\max_{x \in |x - \hat{x}|} |E(x; \hat{x}, y, z)| \leq \Lambda_{k} - x^{+} C \left( \max_{1 \leq k \leq n} \left| \frac{y_{k} - y_{k-1}}{\hat{x}_{k} - \hat{x}_{k-1}} \right| \right) b_{n}(C, z), \quad (44)$$

for \(C\) in \[39\] and

$$b_{n}(C, z) := \max_{1 \leq k \leq n} \sum_{j=1}^{n} (\hat{x}_{j} - \hat{x}_{j-1}) \left| \sum_{j=i}^{n} a_{k,j}(C, z) \right|. \quad (45)$$

When \(C\) has \(O(n)\) elements we can evaluate \(b_{n}(C, z)\) in \(O(n^{2})\) operations. In fact, we can compute all the weights \(\lambda_{k}(\hat{x})\) and \(z\) in \(O(n^{2})\) operations. Next, we can evaluate \(P(c_{k}; \hat{x}, z)\) and \(\ell_{k}(c_{k}; \hat{x})\), before obtaining all the \(a_{k,j}\) in \[42\] for a given \(k\) in \(O(n)\) operations using the identity

$$\ell_{j}(x; \hat{x}) = \frac{\lambda_{j}(\hat{x}) x - \hat{x}_{j}}{\lambda_{k}(\hat{x}) x - \hat{x}_{j}} \ell_{k}(x; \hat{x}). \quad (46)$$
Theorem 6

Under the conditions (10)–(13), let \( f \) be a function with Lipschitz constant \( L \) and consider the barycentric formula \( q \) in (3) satisfies

\[
|q(\hat{x}; \hat{\mathbf{z}}, \hat{\lambda}(\mathbf{x})) − P(\hat{x}; \hat{\mathbf{z}}, \mathbf{y})| ≤ 1.3 \times 10^{-15} L n (2.9 + \log n)^2.
\]

For each \( n \), set \( C \) and vector \( z \), we obtain a single number \( b_n(C, z) \), which we can compute off-line. In Salzer’s case, our computations with \( C \) formed by the rounded Chebyshev points corresponding to \( 2n \) lead to the \( b_n(C, z) \) in Table 3.

We now present a bound for the error in Step II for the second formula based only on the \( z_k \). This bound does not take the cancelation of rounding errors into account and, consequently, it is worse than the bound obtained by combining Table 3, equation (44) and Theorem 4.

Theorem 5

Let \( f : [\hat{x}^-, \hat{x}^+] \rightarrow \mathbb{R} \) be a function with Lipschitz constant \( L \) such that \( y_k = f(\hat{x}_k) \). If \( \hat{x} \in [\hat{x}^-, \hat{x}^+] \) then

\[
|E(\hat{x}, \hat{\mathbf{z}}, \mathbf{y}, z)| \leq L \tau(\hat{\mathbf{z}}) ||z||_1 + \Lambda x_{\hat{x}_- \hat{x}_+} ||z||_\infty \sup_{s \in [\hat{x}^-, \hat{x}^+]} |f(s) − P(s, \hat{\mathbf{z}}, \mathbf{y})|
\]

\[
\leq L \left( \tau(\hat{\mathbf{z}}) ||z||_1 + \Lambda x_{\hat{x}_- \hat{x}_+} (1 + \Lambda x_{\hat{x}_- \hat{x}_+}) ||z||_\infty \frac{(\hat{x}^+ − \hat{x}^-) \pi}{4(n + 1)} \right).
\]

where

\[
\tau(\hat{\mathbf{z}}) := \max_{\hat{x} \in [\hat{x}^-, \hat{x}^+]} |f_\hat{x}(\hat{\mathbf{z}}, \hat{x} − \hat{x}_k)|. \tag{49}
\]

The next lemma shows that the term \( \tau(\hat{\mathbf{z}}) \) in Theorem 5 is not much different from \( \tau(x) \) when \( \Delta \) and \( ||z||_\infty \) are small:

Lemma 7

Under the conditions (10)–(13), consider \( \Delta \) in (16) and \( z = z(\hat{\mathbf{z}}, \hat{\lambda}(\mathbf{x})) \) in (17). If \( \Delta < 1 \) and \( ||z||_\infty < 1 \) then

\[
\tau(\hat{\mathbf{z}}) \leq \frac{\tau(x)}{(1 − \Delta)(1 − ||z||_\infty)}.
\]

The bounds above and the results in (16) lead to the following theorem:

Theorem 6

Under the conditions (10)–(13), let \( f : \mathbb{R} \rightarrow \mathbb{R} \) be a function with Lipschitz constant \( L \) and consider \( y = f(\hat{x}) \). In Salzer’s case, for all \( x \in [−1, 1] \), the second barycentric formula \( q \) in (3) satisfies

\[
|q(\hat{x}; \hat{\mathbf{z}}, \hat{\lambda}(\mathbf{x})) − P(\hat{x}; \hat{\mathbf{z}}, \mathbf{y})| ≤ 1.3 \times 10^{-15} L n (2.9 + \log n)^2.
\]

| \( n \) | \( b_n(C, z) \) | \( n \) | \( b_n(C, z) \) | \( n \) | \( b_n(C, z) \) | \( n \) | \( b_n(C, z) \) |
|---|---|---|---|---|---|---|---|
| 10 | 5.2e-17 | 100 | 2.1e-16 | 1.000 | 4.5e-16 | 10.000 | 5.5e-16 | 100.000 | 9.0e-16 |
| 20 | 1.2e-16 | 200 | 2.9e-16 | 2.000 | 5.0e-16 | 20.000 | 8.6e-16 | 200.000 | 8.7e-16 |
| 30 | 1.4e-16 | 300 | 3.6e-16 | 3.000 | 4.7e-16 | 30.000 | 7.1e-16 | 300.000 | 1.3e-15 |
| 40 | 1.1e-16 | 400 | 3.0e-16 | 4.000 | 5.3e-16 | 40.000 | 8.1e-16 | 400.000 | 1.1e-15 |
| 50 | 1.7e-16 | 500 | 3.4e-16 | 5.000 | 5.7e-16 | 50.000 | 8.4e-16 | 500.000 | 1.2e-15 |
| 60 | 1.7e-16 | 600 | 3.9e-16 | 6.000 | 6.2e-16 | 60.000 | 9.1e-16 | 600.000 | 1.2e-15 |
| 70 | 1.3e-16 | 700 | 3.9e-16 | 7.000 | 6.6e-16 | 70.000 | 7.7e-16 | 700.000 | 1.4e-15 |
| 80 | 1.6e-16 | 800 | 5.6e-16 | 8.000 | 7.1e-16 | 80.000 | 8.0e-16 | 800.000 | 1.5e-15 |
| 90 | 2.0e-16 | 900 | 6.3e-16 | 9.000 | 7.9e-16 | 90.000 | 8.9e-16 | 900.000 | 1.2e-15 |
Moreover, for every polynomial $Q$ with degree at most $n$ and $M$ in (31), we have
\[ |q(\hat{x}; \hat{x}, y, \hat{\lambda}(x^*)) - P(\hat{x}; \hat{x}, y)| \leq 1.3 \times 10^{-15} L n (2.9 + \log n)^2 + M (1.4 \log n + 2). \]

On the other hand, Theorem 5 Lemma 2 and equation (14) lead to the following:

**Theorem 7** Under the conditions (10)–(13), let $f : \mathbb{R} \to \mathbb{R}$ be a function with Lipschitz constant $L$ and consider $y = f(\hat{x})$. In Salzer’s case, with $n \leq 10^9$, and the set $C$ formed by the rounded Chebyshev points corresponding to $2n$, for all $\hat{x} \in [-1, 1]$ the second barycentric formula $q$ in (2) satisfies
\[ |q(\hat{x}; \hat{x}, y, \hat{\lambda}(x^*)) - P(\hat{x}; \hat{x}, y)| \leq 0.8 L b_n(C, z) (2.3 + \log n). \]

Moreover, for every polynomial $Q$ with degree at most $n$ and $M$ in (31), we have
\[ |q(\hat{x}; \hat{x}, y, \hat{\lambda}(x^*)) - P(\hat{x}; \hat{x}, y)| \leq 0.8 L (b_n(C, z) + 3.7 \times 10^{-16}) (2.3 + \log n) + M (1.4 \log n + 2). \]

In order to use Theorem 7, we may need to rely on empirical data regarding $b_n$, like the data presented in Table 3. However, Theorem 7 yields a sharper upper bound than Theorem 6 for large $n$. For instance, we have $b_n \leq 2 \times 10^{-15}$ in all entries in Table 3 and no experiment we performed resulted in $b_n$ greater than $3 \times 10^{-15}$. Therefore, our empirical data in combination with Theorem 7 suggests that the forward error in Step II for the second formula will be at most $2.2 \times 10^{-15} L (2.3 + \log n)$ for $n$ up to one million, and this bound is much smaller than the one provided by Theorem 6.

Theorems 6 and 7 bound the forward error in Step II for the second formula, and the next theorem presents a bound on the forward error on Step III. By combining these bounds with the bounds above we obtain an overall bound for the numerical forward errors for the second formula.

**Theorem 8** Under the conditions (10)–(13), let $\varepsilon$ be the machine precision and assume that $(2n + 5) \varepsilon < 1$ and $\|\xi(\hat{x}(\hat{\lambda}(\hat{x})), \hat{\lambda})\|_{\infty} (1 + A_{\hat{x}^-}, \hat{x}^+, \hat{\lambda}) < 1$ and define
\[ \Lambda := \frac{A_{\hat{x}^-}, \hat{x}^+, \hat{\lambda}}{1 - \|\xi(\hat{x}(\hat{\lambda}(\hat{x})), \hat{\lambda})\|_{\infty} (1 + A_{\hat{x}^-}, \hat{x}^+, \hat{\lambda})}. \]

If $(n + 2)(2 + \Lambda) \varepsilon < 1$ then for every $\hat{x} \in [\hat{x}^-, \hat{x}^+]$ and $y \in \mathbb{R}^{n+1}$ the computed value $\text{fl}(q(\hat{x}; \hat{x}, y, \hat{\lambda}))$ satisfies
\[ \|\text{fl}(q(\hat{x}; \hat{x}, y, \hat{\lambda})) - q(\hat{x}; \hat{x}, y, \hat{\lambda})\|_{\infty} \varepsilon \leq \Lambda \theta + 2n + 5 \frac{\varepsilon}{1 - (2n + 5) \varepsilon} \|y\|_{\infty} \varepsilon, \]

for
\[ \theta := \frac{(1 + \Lambda) (n + 2)}{1 - (n + 2)(2 + \Lambda) \varepsilon}. \]

In Salzer’s case with $\varepsilon \leq 2.3 \times 10^{-16}$, the forward error in Step III is bounded by
\[ \|\text{fl}(q(\hat{x}; \hat{x}, y, \hat{\lambda}(x^*)) ) - q(\hat{x}; \hat{x}, y, \hat{\lambda}(x^*))\|_{\infty} \varepsilon \leq 0.26 (2n + 5) (1.6 + \log n) (5.8 + \log n) \|y\|_{\infty} \varepsilon \approx 0.5 n \log^2 (n) \|y\|_{\infty} \varepsilon. \]

Finally, we note that (12) also presents bounds for Step III applicable in Salzer’s case. On the other hand, some bounds in that article involve the Lipschitz constant, on the other hand, they do not have $O(\log n)$ factors.
7 Improving the accuracy of the first formula in Salzer’s case

The results in the previous sections show that the accuracy of both barycentric formulae in the Step II in Figure 1 is affected by the relative errors in the length of the intervals \([x_j, x_k]\), which we measure by \(\delta_k, \Delta\) and lead to the relative errors \(z_k\) in the weights. Moreover, Table 1 and Figure 2 show that the first formula is quite sensitive to errors in this second step.

In Salzer’s case, when \(n\) is large, the \(\delta_k(k+1)\) for \(k\) near 0 are much larger than the \(\delta_k(k+1)\) for \(k\) near \(n/2\), because the intervals \([x_k, x_{k+1}]\) for \(k\) small have length of order \(1/n^2\) whereas the intervals \([x_k, x_{k+1}]\) for \(k\) near \(n/2\) have length of order \(1/n\). As a result, the \(z_k\)s tend also to be larger for \(x_k\) near \(\pm 1\). The same discrepancy in the sizes of the \(\delta_{jk}\) and the \(z_k\) will happen whenever we round the nodes as usual and they cluster around a point \(c \in [x^-, x^+]\). This section proposes an efficient way to reduce the largest \(z_k\), by improving the accuracy of the nodes near the points at which they cluster, and presents experiments showing that this procedure is effective for the first formula in Salzer’s case. In principle, our strategy would work for any family of nodes, and any accumulation point \(c\), but it is particularly appropriate for the Chebyshev points, because in this case the trigonometric identity

\[
1 - \cos(k\pi/n) = 2\sin^2(k\pi/(2n))
\]

allow us to evaluate the difference \(x_k - x_0\) accurately in double precision, and the difference \(x - x_0 = 1 + x\) is computed exactly for a floating point number \(x \in [-1, -1/2]\). In other situations, for a general cluster point \(c\), the implementation of our strategy may be more difficult.

This section is divided into three subsections. In the first subsection, we describe a new finite precision representation of the nodes and we explain why it improves the accuracy of the first formula. We then present experimental results showing that our finite precision representation of the nodes is practical in terms of performance and accuracy. We conclude with some remarks about the new node representation.

7.1 A new finite precision representation for the nodes

We now describe a finite precision representation of the nodes that lead to smaller \(\delta_{jk}\) and \(z_k\) without the use of quadruple precision. We partition the interval \([x^-, x^+]\) into sub-intervals which we will refer to as bins. The \(l\)-th bin has a base point \(b_l\) and the nodes \(x_k\) in this bin are represented using \(r_k := x_k - b_l\), so that \(x_k = b_l + r_k\). We store rounded versions \(\hat{r}_k\) of \(r_k\) instead of \(\hat{x}_k\). This idea is similar to Dekker’s [5] but our approach is more economical since we only store one \(b_l\) per bin. For example, we could use the bins \([-1, -0.5), [-0.5, 0.5]\) and \((0.5, 1]\) as illustrated in Figure 6.

![Fig. 6: Partitioning \([-1, 1]\) into bins with base points \(b_0 = -1, b_1 = 0\) and \(b_2 = 1\).](image)

\[
\begin{array}{cccccc}
\text{bin 0} & \text{bin 1} & \text{bin 1} & \text{bin 2} \\
\hline
b_0 = -1 & b_1 = 0 & b_1 = 0 & b_2 = 1 \\
-1/2 & 1/2 & & \\
r_k = 2\sin^2(\frac{2k\pi}{n}) & r_k = \sin(\frac{2k\pi}{n}) & r_k = 2\sin^2(\frac{2k\pi}{n}) & \\
\end{array}
\]
The effects of rounding errors in the nodes on barycentric interpolation

The $r_k$ in bins 0 and 2 can be computed with much smaller absolute errors than the corresponding $\hat{x}_k$. The relative errors in $\hat{r}_k$ and $\hat{x}_k$ have the same order of magnitude. However, $\hat{r}_k$ is smaller than $\hat{x}_k$ and the absolute error in $\hat{r}_k$ for $x_k$ near the border is orders of magnitude smaller than the error in $\hat{x}_k$ for large $n$.

The key point in the representation $x_k = b_l + r_k$ is the possibility of computing the differences $x - x_k$ without evaluating $x_k$ explicitly. Instead, given $x \in [\hat{x}^-, \hat{x}^+]$,

(a) we find $l$ such that $x$ is in the $l$-th bin and define $r_x := x - b_l$.

(b) Given a node $x_k$ in the $m$-th bin, we evaluate $x - x_k$ as $(b_l - b_m) + (r_x - r_k)$.

The steps (a) and (b) are accurate because:

(i) We choose the bins and their bases so that we can apply the following lemma and conclude that there is no rounding errors in the evaluation of $r_x$ in step (a).

**Lemma 8 (Sterbenz’s Lemma)** If subtraction is performed with a guard digit and $x/2 \leq y \leq 2x$ then $x - y$ is computed exactly.

(ii) We choose base points $b_l$ such that the differences $b_l - b_m$ are computed exactly.

(iii) When $x$ and $x_k$ are very close, the difference $b_l - b_m$ is small and the difference in step (b) is computed with high relative accuracy.

The statements above can be formalized as follows:

**Theorem 9** In Salzer’s case, with $x_k = b_l + r_k$ for $r_k$ as in Figure 2 consider the nodes $\tilde{x}_k = b_l + \hat{r}_k$. If $\hat{r}_k = r_k(1 + \theta_k)$, with $|\theta_k| \leq 4.6 \times 10^{-16}$, then the numbers $\delta_{jk}(\tilde{x}, x)$ in (14) satisfy

$$\sum_{j \neq k} |\delta_{jk}(\tilde{x}, x)| \leq \|\theta\|_{\infty} (3.2 + 2.3n + 4.3n \log(n + 1)) \quad \text{ (56)}$$

The bound (56) is of order $O(\varepsilon n \log n)$ whereas the analogous bound for nodes rounded as usual $z$ is of order $\Theta(\varepsilon n^2)$. This explains why bins improve the accuracy of the first formula in Salzer’s case. Moreover, the bound (56), in combination with Lemma 1, leads to smaller upper bounds for the forward errors for the second formula in Theorem 4 and to smaller bounds on the backward errors for the second formula presented in [13].

7.2 Nodes in bins are practical

In this section, we compare common implementations for both barycentric formulae with implementations based on bins (as described in Appendix B). Our results are presented in tables 4, 5 and 6. Table 4 shows that by using bins we reduce the errors introduced by Step II for the first formula. However, even with this reduction, the errors introduced by Step II for the first formula are larger than the same errors for the second formula. Table 5 also shows that using bins did not improve the second formula, because the Step II errors with usual nodes for this formula are already small.
Table 4: Maximum error in Step II for $f(x) = \cos(100x)$ and $x \in X_{-1,n}$.

| $n + 1$ | First Formula | Second Formula |
|---------|---------------|----------------|
|         | as usual      | 3 bins 39 bins 79 bins | as usual      | 3 bins 39 bins 79 bins |
| $10^3$  | 8.9e-12       | 5.0e-15 2.1e-15 2.0e-15 | 4.4e-16       | 9.7e-17 6.8e-17 6.8e-17 |
| $10^4$  | 4.4e-10       | 7.8e-15 9.0e-15 2.7e-15 | 9.0e-17       | 9.0e-17 9.0e-17 9.0e-17 |
| $10^5$  | 7.1e-08       | 8.5e-15 7.8e-15 3.1e-15 | 8.0e-17       | 7.7e-17 7.7e-17 7.7e-17 |
| $10^6$  | 2.1e-06       | 1.1e-14 1.6e-14 1.2e-14 | 7.2e-17       | 7.7e-17 7.7e-17 7.7e-17 |

Table 5 considers the error in the three steps in Figure 1 and shows that, indeed, the use of bins makes the first formula as accurate as the second overall. Table 6 shows that, in these experiments, nodes in bins lead to the same performance as the usual ones. These results can be explained by the need for only one extra sum per node in bin and that the dominant factors in performance are access to memory and divisions. In order to evaluate $\Delta x_k := x - x_k$ for all $x_k$ in the $m$-th bin when $x$ is in the $l$-th bin, we compute $\Delta b = b_l - b_m$ only once and then compute $\Delta x_k = \Delta b + (r_x - r_k)$. The cost incurred in evaluating this expression is less than twice the cost incurred in evaluating $\Delta x_k$ as usual because $\Delta b$ stays in the cache.

Table 5: Maximum overall error for $f(x) = \cos(100x)$ and $x \in X_{-1,n}$.

| $n + 1$ | First Formula | Second Formula |
|---------|---------------|----------------|
|         | as usual      | 3 bins 39 bins 79 bins | as usual      | 3 bins 39 bins 79 bins |
| $10^3$  | 8.9e-12       | 1.2e-14 1.2e-14 1.2e-14 | 1.1e-14       | 1.1e-14 1.0e-14 1.0e-14 |
| $10^4$  | 4.4e-10       | 3.5e-14 3.3e-14 3.1e-14 | 3.0e-14       | 3.0e-14 3.0e-14 3.0e-14 |
| $10^5$  | 7.1e-08       | 8.9e-14 9.6e-14 9.6e-14 | 8.9e-14       | 8.7e-14 8.7e-14 8.7e-14 |
| $10^6$  | 2.1e-06       | 2.4e-13 2.5e-13 2.4e-13 | 1.7e-13       | 1.6e-13 1.6e-13 1.6e-13 |

Table 6: Normalized times (usual first formula = 1).

| $n + 1$ | First Formula | Second Formula |
|---------|---------------|----------------|
|         | as usual      | 3 bins 39 bins 79 bins | as usual      | 3 bins 39 bins 79 bins |
| $10^3$  | 1.00          | 1.01 1.05 1.07 | 1.40          | 1.40 1.40 1.40 |
| $10^4$  | 1.00          | 1.00 1.01 1.01 | 1.40          | 1.41 1.41 1.41 |
| $10^5$  | 1.00          | 0.99 0.99 0.99 | 1.40          | 1.39 1.39 1.39 |
| $10^6$  | 1.00          | 1.00 1.00 1.00 | 1.38          | 1.38 1.37 1.38 |

Tables 5 and 6 indicate that, in these experiments, the first formula with bins is competitive in terms of performance and accuracy, whereas the accuracy of the usual first formula is unacceptable. We doubt that the accuracy will be affected significantly with the use of other compilers, machines and programming languages. The performance, on the other hand, depends on the machine, the compiler and the language. For instance, Fortran sometimes leads to faster code than C++, and by using a dif-
ferent language we could obtain a different relation among the performance of the several alternatives described in the previous tables.

7.3 Final remarks

The approximate nodes \( \tilde{x}_k \) in Theorem 9 may not be floating point numbers, in the same way that Dekker’s numbers usually are not floating point numbers. This is not a problem if we already have the \( y_k \). In this case, we do not need the \( \tilde{x}_k \), because both formulae can be expressed in terms of the \( y_k \) and the differences \( x - \tilde{x}_k \), and such differences can be computed without the explicit value of \( \tilde{x}_k \). However, any \( \tilde{x}_k \) that is not a floating point number complicates the evaluation of \( y_k = f(\tilde{x}_k) \). We could handle this problem in three ways:

- Evaluate \( f(\hat{x}_k) \), where \( \hat{x}_k \) is the floating point number closest to \( \tilde{x}_k \).
- Estimate \( f(\tilde{x}_k) \) as \( f(r_k) + f'(\tilde{x}_k)(\tilde{x}_k - \hat{x}_k) \).
- Evaluate \( f(\tilde{x}_k) \), on-line or off-line, using higher precision arithmetic.

In any case, the readers will need to take these considerations into account should they decide to apply the ideas presented in this section.

There are many choices for the bins and their base points but there isn’t a single choice that is optimal for all compilers, processors and instruction sets. For instance, the advances in hardware may lead to efficient combinations of integer and floating point arithmetic. We experimented with \( \hat{x}_k \) represented as 64 bit integers and the resulting code was 50% slower than the one using only floating point arithmetic. However, with integer arithmetic we reduced the Step II errors by a factor of \( 10^3 \), due to a better use of the 11 bits that represent the exponent in IEEE754 double precision.

The errors in Step II for the implementations described in the previous subsection are much smaller than the Step III errors. As a result, the overall error is determined by Step III and the gain in accuracy in Step II due to the use of integer arithmetic was irrelevant. Therefore, our mixing of floating point and integer arithmetic is not competitive at this time. However, parallel usage of the hardware dedicated to floating point and integer arithmetic could change this.

Finally, in extreme situations, the use of bins can also improve the accuracy of the second formula. In our experiments with \( f(x) = \cos(10^5x) \) and one million nodes, the maximum error with the usual second formula was of the order \( 10^{-12} \) and the error with the second formula with nodes in bins was of the order \( 10^{-13} \). Moreover, the results in section 3 of [13] indicate that for some functions with large Lipschitz constants the backward error for the second formula could also be reduced by the use of bins.

A Proofs

This appendix proves the results stated in the previous sections. We state three more lemmas, after that we prove all lemmas in the order in which they were stated, we then prove the theorems, also in the order in which they were stated.
Lemma 9 Given a vector \( v = (v_1, v_2, \ldots, v_k)^T \in \mathbb{R}^{n+1} \), define \( s := \sum_{k=0}^{n} v_k \), \( a := \sum_{k=0}^{n} |v_k| \).

\[
x_- := x_-(v) := \sum_{k=0}^{n} \min \{ v_k, 0 \} \quad \text{and} \quad x_+ := x_+(v) := \sum_{k=0}^{n} \max \{ v_k, 0 \}.
\]

If \( s_- < 1 \) then

\[
-1 \leq \frac{x_-}{1-x_-} \leq \frac{x_+}{1+x_+} \leq \frac{x_-}{1-s_-} \quad (57)
\]

and

\[
\prod_{k=0}^{n} (1 + v_k) - 1 - s \geq s x_- \geq -\frac{s^2}{4} \quad (58)
\]

Moreover, if \( s_- < 1 \) and \( s < 1 \) then

\[
\prod_{k=0}^{n} (1 + v_k) - 1 - s \leq \frac{s^2}{1-s} \quad (59)
\]

Finally, if \( a < 1 \), then, for \( 0 \leq m \leq n \), the product

\[
P := \left( \prod_{k=0}^{m} \frac{1}{1+v_k} \right) \left \{ \prod_{k=n+1}^{n+m+1} (1+v_k) \right \}
\]

satisfies

\[
1 - a \leq P \leq \frac{1}{1-a} \quad \text{and} \quad |P - 1| \leq \frac{a}{1-a} \quad (60)
\]

Lemma 10 For \( 1 \leq k \leq n/2 \), the Chebyshev points of the second kind satisfy

\[
\sum_{j=0}^{k-1} \frac{1}{x_k - x_j} \leq \frac{1}{2 \sin \left( \frac{\pi}{n} \right) \sin \left( \frac{\pi}{2} \right)} (1 + k \log(4k - 1)) \leq \frac{n^2}{4 \sqrt{2} k^2} (1 + k \log(4k - 1)) \quad (61)
\]

and

\[
\sum_{j=1}^{k} \frac{1}{x_k - x_j} \leq \frac{3n^2}{4 \sqrt{2} k \log(4k + 1)} \quad (62)
\]

In particular, since \( \frac{61}{62} \) and \( \frac{62}{63} \) decrease with \( k \), for \( 1 \leq k \leq n/2 \)

\[
\sum_{j=0}^{k} \frac{1}{x_k - x_j} \leq \frac{n^2}{4 \sqrt{2}} (1 + \log 3 + 3 \log 5) < 1.2246 n^2 \quad \text{and} \quad \sum_{j=1}^{k} \frac{1}{x_k - x_0} \leq \frac{n^2}{12} < 0.82247 n^2.
\]

Moreover, if \( n \geq 10 \) then, for \( 0 \leq k \leq n \),

\[
\frac{1}{x_k - x_0} \leq 0.20432 n^2.
\]

Lemma 11 If \( x^- = x_0, x^+ = x_n \) and

\[
\mu := \min_{0 \leq j \leq n} (x_{k+1} - x_k)
\]

is such that \( 2\mu \| k - x \|_\infty < 1 \), then \( \delta_k \), \( \Delta_k \) in \( \frac{43}{44} \) and \( a_k = \sum_{j \neq k} \delta_k \) satisfy

\[
|\delta_k| \leq \frac{\kappa}{x_j - x_k} \quad \text{for} \quad j \neq k \quad \text{and} \quad a_k \leq \kappa \sum_{j \neq k} \frac{1}{x_j - x_k} \quad (63)
\]

and, for \( 0 \leq k \leq n \),

\[
\Delta_k \leq \kappa \left( \sum_{j=0}^{k-1} \frac{1}{x_k - x_j} + \frac{2}{x_k - x_n} + \sum_{j=k+2}^{n} \frac{1}{x_k - x_{j+1}} \right),
\]

for

\[
\kappa := \frac{2 \| k - x \|_\infty}{1 - 2\mu \| k - x \|_\infty} \quad (65)
\]
A.1 Proofs of the lemmas

Proof (Lemma 2) According to definition (3), we have
\[ \lambda_k(x) = \frac{1}{\prod_{j<k} (x_k - x_j)} \quad \text{and} \quad \lambda_d(x) = \frac{1}{\prod_{j<k} (\hat{x}_k - x_j)}. \]

Using definitions (14) and (17) we obtain
\[ z_k = \frac{\lambda_d(x)}{\lambda_d(x_k)} - 1 = \left( \prod_{j<k} \frac{x_k - x_j}{x_k - x_j} \right) - 1 = \left( \prod_{j<k} \frac{1}{1 + \delta_{jk}} \right) - 1, \]
and the bounds on \( z_k \) in Lemma 2 follow from Lemma 9. Similarly,
\[ z_d = \frac{\lambda_d(x)}{\lambda_d(x_k)} - 1 = \left( \prod_{j<k} \frac{x_k - x_j}{x_k - x_j} \right) - 1 = \left( \prod_{j<k} (1 + \delta_{jk}) \right) - 1, \]
and the bound in \( |z_d| \) also follows from Lemma 9. \( \square \)

Proof (Lemma 3) Consider a set of indices \( K = \{0, 1, \ldots, n\} \). If \( \hat{x} = x_k \) for \( k \in K \) then we can take \( x = x_k \) and using the definition (14) we obtain
\[ \frac{x - x_j}{x - x_j} - 1 = |\delta_{jk}| \]
for all \( j \in K \). On the other hand, if \( \hat{x} \notin x \) then Corollary 3 in (13) shows that there exists \( x \in [x^-, x^+] \setminus \{x_0, \ldots, x_n\} \) which satisfies (22) and, for \( 0 \leq j \leq n \),
(a) If \( x^- < \hat{x} < x_k \) then
\[ \frac{x - x_j}{x - x_j} - 1 \leq \max\{ |\delta_{jk}|, |\delta_{jk+1}| \}. \]
(b) If \( k^- \leq k < k^+ \) and \( x_k < \hat{x} < x_{k+1} \) then
\[ \frac{x - x_j}{x - x_j} - 1 \leq \max\{ |\delta_{jk}|, |\delta_{jk+1}| \}. \]
(c) If \( x_k < \hat{x} < x^+ \) then
\[ \frac{x - x_j}{x - x_j} - 1 \leq \max\{ |\delta_{jk^+}|, |\delta_{jk^-}| \}. \]
where \( k^- \) and \( k^+ \) are such that \( x_{k^+} \) is the first node larger than \( \hat{x}^- \) and \( x_{k^-} \) is the last node smaller than \( \hat{x}^+ \). Defining \( d_j = (x - x_j) / (\hat{x} - x_j) - 1 \), we can write the product in the left hand side of (21) as
\[ \prod_{j \in K} (\hat{x} - x_j) \left( \prod_{j \in K} \left( \frac{x - x_j}{x - x_j} \right) \right) = \left( \prod_{j \in K} \frac{1}{1 + d_j} \right) \prod_{j=0}^{n} (x - x_j) = (1 + \beta) \prod_{j=0}^{n} (x - x_j) \]
for
\[ \beta := \prod_{j \in K} \frac{1}{1 + d_j} - 1. \]
In case (a) above we have \( \sum_{j \in K} |d_j| \leq \Delta^- \), in case (b) or when \( \hat{x} = x_k \) for \( k \notin K \) we have \( \sum_{j \in K} |d_j| \leq \Delta_k \) and in case (c) we have \( \sum_{j \in K} |d_j| \leq \Delta^+ \). In all cases Lemma 3 follows from Lemma 9. \( \square \)

Proof (Lemma 4) Recall that we are assuming that \( \|x - x\|_\infty \leq 4.6 \times 10^{-10} \). Lemma 10 shows that
\[ \mu = \frac{1}{\min_j |x_j - x_k|} \leq 0.20432n^2, \] (66)
and the constant $\kappa$ in Lemma 11 is bounded by

$$
\kappa = \frac{2 \|k - \xi\|_\infty}{1 - 2 \|k - \xi\|_\infty \times 0.20432 \times \pi^2} \leq \frac{2 \|k - \xi\|_\infty}{1 - 9.2 \times 10^{-16} \times 0.20432 \times 4 \times 10^4} \leq 2.0016 \|k - \xi\|_\infty.
$$

(67)

The last two equations and (63) show that

$$
\delta \leq \kappa \mu \leq 2.0016 \times 0.20432 \|k - \xi\|_\infty n^2 \leq 0.40897 \|k - \xi\|_\infty n^2 \leq 1.8813 \times 10^{-16} n^2 \leq 7.5252 \times 10^{-4}.
$$

(68)

Lemmas 10 and 11 and (66) and (67) lead to the following bound on $a_k = \sum_{j=1}^k \mid \delta_j \mid$:

$$
a_k \leq \sum_{j=1}^k \frac{\kappa}{\mid x_j - x_k \mid} \leq 2.4512 \|k - \xi\|_\infty n^2 \leq 1.1276 \times 10^{-13} n^2 \leq 4.5104 \times 10^{-3}.
$$

(69)

Lemma 1 and the bound on $\alpha_k$ above yield

$$
\mid a_k \mid \leq \frac{\alpha_k}{1 - \alpha_k} \leq \frac{2.4512 \|k - \xi\|_\infty n^2}{1 - 4.5104 \times 10^{-3}} \leq 2.4624 \|k - \xi\|_\infty n^2 \leq 1.1328 \times 10^{-15} n^2 \leq 4.5312 \times 10^{-3}.
$$

(70)

The same bounds hold for $\zeta_k$, because Lemma 1 also states that $\mid \zeta_k \mid \leq \alpha_k/(1 - \alpha_k)$. Lemmas 10 and 11 also lead to

$$
\Delta_0 \leq \kappa \left( \frac{2}{x_1 - x_0} + \sum_{j=2}^k \frac{1}{x_j - x_1} \right) \leq \kappa \left( 2 \times 0.20432 \times n^2 + \frac{3n^2}{4\sqrt{2}} \log 5 \right) \leq 2.5264 \|k - \xi\|_\infty n^2.
$$

For $1 \leq k \leq n/2$, noting that the second sum in (64) starts at $k + 2$, we obtain

$$
\Delta_k \leq \frac{2\kappa \mu + \frac{\kappa n^2}{4\sqrt{2}}}{x_{k+1} - x_k} \left( 1 + \log 3 + \frac{3}{2} \log 9 \right) \leq 2.7267 \|k - \xi\|_\infty n^2 \leq 1.2543 \times 10^{-15} n^2 \leq 5.0172 \times 10^{-3}.
$$

(71)

Let us now prove the bound on $\|x\|_1$. When $1 \leq k \leq n/2$, Lemmas 10 and 11 show that

$$
a_k \leq \frac{\kappa n^2}{4\sqrt{2}} \psi_k,
$$

(72)

for

$$
\psi_k := \frac{1}{k^2} + \frac{\log (4k - 1)}{k} + \frac{3 \log (4k + 1)}{k}.
$$

When $|x| < 1$, $\log(1 - x) + \log(1 + x) = \log(1 - x^2) < 0$ and $\log(1 + x) \leq x$. Therefore,

$$
\log(1 - 1/4k) + 3\log(1 + 1/4k) \leq \frac{1}{2k} \quad \text{and} \quad \psi_k \leq f(k) := \frac{3}{2k} + \frac{8 \log 2 + 2 \log k}{k}.
$$

(73)

The function $f$ has derivative

$$
f'(x) := -\frac{3}{x^2} + \frac{8 \log 2}{x^2} - \frac{2 \log x}{x^2},
$$

which is negative for $x \geq 1$. Therefore,

$$
\sum_{k=1}^{[\sqrt{n}/2]} \psi_k \leq \int_1^{\sqrt{n}/2} f(x) \, dx = \left. -\frac{3}{2x} + 8 \log 2 \log x + \log^2 x \right|_{x=1}^{x=n/2}
$$

$$
= -\frac{3}{n} + 8 \log 2 (\log n - \log 2) + (\log n - \log 2)^2 + \frac{3}{2} \leq \log n (\log n + 6 \log 2) + \frac{3}{2} - 7 \log^2 2.
$$

(74)
By symmetry, and equations (72) and (74), we have

$$\sum_{k=0}^{n} a_k \leq 2 (a_0 + a_1) + \frac{K n^2}{2 \sqrt{2}} \sum_{k=0}^{n/2} \theta_k \leq A \|x - \hat{x}\|_n^2,$$

for

$$A := \frac{2 (a_0 + a_1)}{\|x - \hat{x}\|_n^2} + \frac{1.0008}{\sqrt{2}} \left( \log n (\log n + 6 \log 2) + \frac{3}{2} - 7 \log^2 2 \right).$$

Using bound (69), the identity \((2.9 + \log n)^2 = \log^2 n + 2 \times 2.9 \times \log n + 2.9^2\), hypothesis \(n \geq 10\) and the fact that \(6 \log 2 - 2 \times 2.9 < 0\), we deduce that

$$A \leq \frac{1.0008}{\sqrt{2}} (2.9 + \log n)^2 + \frac{1.0008}{\sqrt{2}} \left( \frac{3}{2} - 7 \log^2 2 - 2.9^2 + (6 \log 2 - 2 \times 2.9) \times 10 \right) + \frac{2 (a_0 + a_1)}{\|x - \hat{x}\|_n^2} \leq \frac{1.0008}{\sqrt{2}} (2.9 + \log n)^2 - 0.13941 < 0.70768 (2.9 + \log n)^2.$$

Moreover, \(21\) and (69) show that \(\|z\| \leq a_i/1 - a_i \leq 1.0046 a_i\). This inequality, the bound on \(A\) above, and (75) yield

$$\|z\| \leq 0.71094 \|x - \hat{x}\|_n (2.9 + \log n)^2 n^2 \leq 3.2704 \times 10^{-16} (2.9 + \log n)^2 n^2 \leq 0.39646$$

and we are done. \(\Box\)

**Proof (Lemma 4)** In the notation of (13), we write the \(\Lambda_{x_0 - x, \hat{x}}\) in (13) as \(\Lambda_{x_0 - x, \hat{x}(x)}\). Moreover, if we take \(w = \hat{\lambda}(x)\) and \(w = \hat{\lambda}(\hat{\lambda})\) then (17) implies that \(\xi = \xi(w, \hat{\lambda}) = \hat{\lambda}(x, \hat{\lambda}(x))\) and the hypothesis (25) shows that \(d = \delta\) and \(\zeta\) satisfy the hypothesis in Theorem 3 in (13). Therefore, (26) follows from Corollary 3 and Theorem 3 in that article. Theorem 1 in (8) combined with the results in (14) yield that, for the Euler-Mascheroni constant \(\gamma < 0.577215665\) and \(n \geq 10\),

$$\Lambda_{x_0, z} \leq \frac{2}{\pi} \left( \log n + \gamma + \log \frac{8}{\pi} + \frac{\pi^2}{144n} \right) \leq 0.63662 (\log n + 1.5127).$$

Moreover, evaluating the term after the first \(\leq\) in the expression above for \(n = 2 \times 10^6\) we obtain

$$\Lambda_{x_0, z} \leq 10.1991.$$ (78)

The bounds on \(d\) and \(z\) in Lemma 4 and (26) lead to

$$\Lambda_{x_0, z} \leq 10.629 \Lambda_{x_0, z}.$$ (79)

This bound and (72) lead to \(\Lambda_{x_0, z} \leq 0.67667 \log n + 1.0326\), and (78) and (79) show that \(\Lambda_{x_0, z} \leq 10.841\).

**Proof (Lemma 5)** Lemma 3 yields \(\beta\) satisfying (28) and equation (29) follows from Lemma 3 and the bound on \(\|\hat{\lambda}_0(x - x)\|_n\) presented in (13).

**Proof (Lemma 6)** Equation (45) yields

$$\sum_{j=0}^{n} a_{k,j} - \frac{a_j}{\|x - \hat{x}\|_n^2} P(c_k; \hat{x}, \hat{z}) - \frac{c_j}{\|x - \hat{x}\|_n^2} P(c_k; \hat{x}, \hat{z}) = 0.$$ (80)

Consider now \(u_j := y_j - y_{j-1}\). It follows that \(y_j = y_0 + \sum_{i=1}^{j} u_i\). Using (80), we obtain

$$h_i(y) = \sum_{j=0}^{n} a_{k,j} y_j = a_i y_0 + \sum_{j=0}^{n} a_{k,j} \left( y_0 + \sum_{i=1}^{n} u_i \right) = \sum_{j=0}^{n} a_{k,j} y_0 + \sum_{j=0}^{n} a_{k,j} u_i.$$ (81)

The constraints (43) imply that \(|u_i| \leq L(\hat{x}_i - \hat{x}_{i-1})\). As a result, (81) leads to

$$|E(c_k; \hat{x}, y, z)| \leq \sum_{i=1}^{n} |h_i(y)| \leq \sum_{i=1}^{n} \sum_{j=0}^{n} a_{k,j} \left| y_0 \sum_{j=0}^{n} a_{k,j} \right| \leq L \sum_{i=1}^{n} \left( \hat{x}_i - \hat{x}_{i-1} \right) \sum_{j=0}^{n} a_{k,j}$$

and we are done. \(\Box\)
Lemma 2 lead to

Proof (Lemma 7) We show that for every \( \tilde{x} \in [\tilde{x}^-, \tilde{x}^+] \) there exists \( x \in [x^-, x^+] \) such that

\[
|\xi(k; x, \tilde{x})(x - \tilde{x})| \leq \frac{|\xi_k(x, \tilde{x})(x - \tilde{x})|}{(1 - |\tilde{x}|^2)(1 - \Delta)}.
\] (82)

When \( \tilde{x} = \tilde{x}_k \) we can satisfy (82) by taking \( x = x_k \). For \( \tilde{x} \notin \{\tilde{x}_0, \tilde{x}_1, \ldots, \tilde{x}_n\} \), equations (17) and (20) and Lemma 3 lead to

\[
|\xi(k; x, \tilde{x})(x - \tilde{x})| \leq \alpha_k(x, \tilde{x}) \sum_{k=0}^{n} |\tilde{x} - \tilde{x}_k| = 1 + \beta \sum_{k=0}^{n} |\tilde{x} - \tilde{x}_k| = 1 + \beta \sum_{k=0}^{n} |x - x_k|
\]

with \( \beta \) such that \( |\beta| \leq \Delta/(1 - \Delta) \). Equation (82) follows from this equation and we are done. \( \square \)

Proof (Lemma 8) Lemma 3 is Theorem 11 in page 229 of [7].

Proof (Lemma 9) If \( v_j v_k \geq 0 \) then \( \tilde{v}_j = v_j + v_k \) and \( \tilde{v}_k = 0 \) satisfy \( |\tilde{v}_j| + |\tilde{v}_k| = |v_j| + |v_k| \) and

\[
(1 + \tilde{v}_j)(1 + \tilde{v}_k) = 1 + v_j + v_k < 1 + v_j + v_k + v_j v_k = (1 + v_j)(1 + v_k).
\]

Therefore, by replacing \( v_j \) and \( v_k \) by \( \tilde{v}_j \) and \( \tilde{v}_k \) we do not change the sums \( s_-(v) \) and \( s_+(v) \), and we decrease the product \( \prod_{k=0}^{n} (1 + v_k) \), because the hypothesis \( s_- < 1 \) implies that all its factors are positive. Applying this argument while there are pairs \( v_j, v_k \) with \( v_j v_k > 0 \), we conclude that

\[
\prod_{k=0}^{n} (1 + v_k) \geq (1 - s_-)(1 + s_+).
\] (83)

This equation leads to

\[
\prod_{k=0}^{n} (1 + v_k) - 1 - s \geq -s_+ s_-.
\]

and the identity \( s_+ s_- = (s^2 - s^4)/4 \) yields the first inequality in (58). Equation (83) also shows that

\[
\prod_{k=0}^{n} 1 + v_k - 1 \leq \frac{1}{(1 - s_-)(1 + s_+)} - 1 = \frac{s_-}{s_+},
\]

and this proves the second inequality in (57). The set \( C := \{v \in \mathbb{R}^{n+1} \text{ with } s_-(v) < 1\} \) is convex because the function \( s_- \) is convex. Let us define \( h: C \to \mathbb{R} \) by

\[
h(v) := \prod_{j=1}^{n} \frac{1}{1 + v_j} - (1 - s).
\]

The function \( h \) has partial derivatives

\[
\frac{\partial h}{\partial v_k}(v) = 1 - \frac{1}{1 + v_k} \prod_{j=1}^{n} \frac{1}{1 + v_j},
\]

\[
\frac{\partial^2 h}{\partial v_k^2}(v) = \frac{2}{(1 + v_k)^2} \prod_{j=1}^{n} \frac{1}{1 + v_j},
\]

and, for \( j \neq k \),

\[
\frac{\partial^2 h}{\partial v_j \partial v_k}(v) = \frac{1}{(1 + v_j)(1 + v_k)} \prod_{j=1}^{n} \frac{1}{1 + v_j}.
\]

Therefore, its Hessian can be written as

\[
\nabla^2 h(v) = \left( \prod_{j=1}^{n} \frac{1}{1 + v_j} \right) D (1 + \mathbb{1} v) D,
\]

where \( D \) is the identity matrix, \( \mathbb{1} \) is the vector with all entries equal to 1 and \( D \) is the diagonal matrix with \( d_{ij} = (1 + v_j)^{-1} \). Therefore, \( \nabla^2 h(v) \) is positive definite and \( h \) is convex. Since \( h(0) = 0 \) and \( \nabla h(0) = 0 \) we have that \( h(v) \geq 0 \) for all \( v \in C \). As a result, we have the first inequality in (57). When \( s < 1 \), this inequality leads to (59).
Finally, when \( a < 1 \), the inequality \( 1 - x \leq 1/(1 + x) \) for \( x \in (-1, 1) \), and the bound (83) lead to
\[
P \geq \prod_{k=1}^{n} (1 - |v_k|) \geq (1 - x_{-}(-|v|)) (1 + x_{+}(-|v|)) = 1 - a
\]
and
\[
P \leq \prod_{k=0}^{n} 1 - |v_k| \leq \frac{1}{(1 - s_{-}(-|v|))(1 + s_{+}(-|v|))} = \frac{1}{1 - a}.
\]
The bound \(|P - 1| \leq a/(1 - a)\) follows from the last two equations and we are done. \(\square\)

**Proof (Lemma 10)** Since \( x_k := -\cos \frac{2\pi k}{n} \), the identity \( \cos a - \cos b = 2\sin \left( \frac{a+b}{2} \right) \sin \left( \frac{a-b}{2} \right) \) leads to
\[
\frac{1}{x_k - x_j} = \frac{1}{2\sin \left( \frac{x_j + x_k}{2} \pi \right) \sin \left( \frac{x_j - x_k}{2} \pi \right)}.
\]
(84)

If \( 0 \leq j < k \), then, by the concavity of the \( \sin \) function in \( [0, k\pi/n] \subset [0, \pi/2] \), we have
\[
\sin \left( \frac{k - j}{2n} \pi \right) \geq \frac{\sin \left( \frac{k\pi}{2n} \right)}{\frac{k\pi}{2n}} \times \frac{k - j}{2n} \pi = \sin \left( \frac{k\pi}{2n} \right) \times \frac{k - j}{k}
\]
and
\[
\sin \left( \frac{k + j}{2n} \pi \right) \geq \frac{\sin \left( \frac{k\pi}{2n} \right)}{\frac{k\pi}{2n}} \times \frac{k + j}{2n} \pi = \sin \left( \frac{k\pi}{2n} \right) \times \frac{k + j}{2k}.
\]
Combining the last two inequalities with the well known inequality \( \sum_{j=1}^{n} \frac{1}{x_j} \leq \log(2m + 1) \) we obtain
\[
\sum_{j=0}^{k-1} \frac{1}{x_k - x_j} \leq \sum_{j=0}^{k-1} \frac{1}{2\sin \left( \frac{x_j + x_k}{2} \pi \right) \sin \left( \frac{x_j - x_k}{2} \pi \right) (k^2 - j^2)}
\]
\[
= \frac{k}{2\sin \left( \frac{k\pi}{2n} \right) \sin \left( \frac{k\pi}{2n} \right)} \sum_{j=0}^{k-1} \left( \frac{1}{k - j} + \frac{1}{k + j} \right)
= \frac{k}{2\sin \left( \frac{k\pi}{2n} \right) \sin \left( \frac{k\pi}{2n} \right)} \sum_{j=0}^{k-1} \left( \frac{1}{k} + \sum_{j=1}^{2k - 1} \frac{1}{j} \right)
\]
This proves the first inequality in (64). The second inequality in (64) follows from the first inequality and the observation that the hypothesis \( 1 \leq k \leq n/2 \) and the concavity of \( \sin \) in \([0, \pi/2] \) yield
\[
\sin \left( \frac{k\pi}{2n} \right) \geq \sin \left( \frac{k\pi}{2n} \right) \geq \frac{2k}{n} \pi \quad \text{and} \quad \sin \left( \frac{k\pi}{2n} \right) \geq \sin \left( \frac{k\pi}{2n} \right) \geq \frac{\sqrt{2}k}{n} \pi.
\]
If \( 0 \leq k < j \leq n \) then \((j - k)/(2n) \pi \leq \pi/2 \) and the concavity of \( \sin x \) in \([0, 3\pi/4] \) yields
\[
\sin \left( \frac{j - k}{2n} \pi \right) \geq \sin \left( \frac{j - k}{2n} \right) \geq \frac{j - k}{2n} \pi \quad \text{and} \quad \sin \left( \frac{j + k}{2n} \pi \right) \geq \sin \left( \frac{j + k}{2n} \right) \geq \frac{\sqrt{2}j + k}{3n} \pi.
\]
If \( k = 0 \), then equations (84) and (85) yield
\[
\sum_{j=0}^{k} \frac{1}{x_j - x_0} = \sum_{j=0}^{k} \frac{1}{2\sin^2 \left( \frac{x_j}{2n} \right)} \leq \sum_{j=0}^{k} \frac{n^2}{2} \pi \sum_{j=0}^{k} \frac{1}{j^2} = \frac{\pi n^2}{12} \leq 0.82247n^2.
\]
When \( 1 \leq k < j \leq n \), equations (84) and the inequalities (85) lead to
\[
\sum_{j=k+1}^{n} \frac{1}{x_j - x_k} \leq \frac{3n^2}{2\sqrt{2} \pi} \sum_{j=k+1}^{n} \frac{1}{j^2 - k^2} = \frac{3n^2}{4\sqrt{2} \pi k} \sum_{j=k+1}^{n} \left( \frac{1}{j-k} - \frac{1}{j+k} \right)
\]
Proof (Theorem 1) A.2 Proofs of the theorems and equation (64) follows from the last four equations. ⊓ ⊔

Finally, by the concavity of \( j \) and when \( \theta \) and equation (19) leads to

\[
\frac{1}{|x_i - x_j|} \leq \frac{1}{|x_1 - x_0|} = \frac{1}{1 - \cos \left( \frac{\pi}{3} \right)} = \frac{1}{2 \sin^2 \left( \frac{\pi}{6} \right)} \leq \frac{1}{2} \left( \frac{\sin \frac{\pi}{6} \frac{\pi}{20} \right)^2 \leq 0.20432n^2,
\]

and this proof is complete. □

**Proof (Lemma 7)** Definition (14) states that \( \delta_k = 0 \) when \( j = k \). For \( 0 \leq j \neq k \leq n \) we have

\[
|\delta_k| = \left| \frac{x_j - x_k}{x_j - x_k} - 1 \right| = \left| \frac{\hat{x}_j - x_k}{x_j - x_k} \right| = \frac{2 \|k - x\|}{\|x_j - x_k\|} \cdot \frac{1}{\|x_j - x_k\|} \leq \frac{\kappa}{\|x_j - x_k\|},
\]

and the first equation in (63) holds. When \( j < k \) we have

\[
\max \left\{ |\delta_k|, |\delta_{(k+1)}| \right\} \leq \frac{\kappa}{x_k - x_j},
\]

when \( j > k + 1 \),

\[
\max \left\{ |\delta_k|, |\delta_{(k+1)}| \right\} \leq \frac{\kappa}{x_j - x_{k+1}},
\]

and when \( j \in \{k, k+1\} \),

\[
\max \left\{ |\delta_k|, |\delta_{(k+1)}| \right\} = |\delta_{(k+1)}| \leq \frac{\kappa}{x_{k+1} - x_k},
\]

and equation (64) follows from the last four equations. □

A.2 Proofs of the theorems

**Proof (Theorem 7)** Note that, for \( 0 \leq k \leq n \),

\[
|Q(x_k) - P(x_k; x, y)| \leq |Q(x_k) - f(x_k)| + |f(x_k) - P(x_k; x, y)| \leq M,
\]

and equation (19) leads to

\[
|Q(x) - P(x; x, y)| \leq A_{x, x^*} \cdot xM \tag{86}
\]

for all \( x \in \mathbb{R}^+ \). Similarly,

\[
|Q(\hat{x}_k) - P(\hat{x}_k; \hat{x}, y)| \leq |Q(\hat{x}_k) - f(\hat{x}_k)| + |f(\hat{x}_k) - f(x_k)| + |f(x_k) - P(x_k; \hat{x}, y)| \leq M + L|\hat{x} - x_k|,
\]

and

\[
|Q(x) - P(x; \hat{x}, y)| \leq A_{x, -x^*} \cdot (M + L\|x - x\|) \tag{87}
\]

Equation (60) follows from (86), (87) and the triangle inequality. Equations (27) and (77) lead to

\[
A_{x, -x^*} \cdot A_{x, -x^*} \leq 2.0629A_{x, -x^*} \leq 1.3133(\log n + 1.5127) \tag{88}
\]

Equation (82) follows from the bound last two bounds and (30). □
Proof (Theorem 2) Given \( \hat{c} \in [\hat{c}^-, \hat{c}^+] \), the arguments used in the proof of Theorem 3.2 in [9] show that the first formula \( p \) with rounded nodes \( \hat{x} \) satisfies

\[
\hat{f}(p(\hat{x}, \hat{y}, \hat{w})) = \sum_{k=0}^{n} \hat{w}_k \gamma_k \frac{(3n+5)k}{x - \hat{x}_k} \prod_{j \neq k} (\hat{x} - \hat{x}_j),
\]

where the \( (m) \) are the Stewart’s relative error counters described in [10]. Lemma 2 yields \( x \) satisfying \( \beta \)

and \( \beta_0, \ldots, \beta_n \) such that \( |\beta| \leq \Delta/(1 - \Delta) \) and

\[
\prod_{j \neq k} (\hat{x} - \hat{x}_j) = (1 + \beta_k) \prod_{j \neq k} (x - x_k).
\]

Combining this equation with \( \epsilon \) we obtain

\[
\hat{f}(p(\hat{x}, \hat{y}, \hat{w})) = \sum_{k=0}^{n} \hat{w}_k \gamma_k \left( 1 + \frac{\hat{w}_k - w_k}{w_k} \right) \prod_{j \neq k} (x - x_k) = p(x, \hat{y}, \hat{w})
\]

for \( \gamma_k \) in \( \epsilon \) and Theorem 2 follows from Lemma 3.1 in [10], which states that \( |(m) - 1| \leq \epsilon (1 - \epsilon) \)

when \( \epsilon < 1 \). \( \square \)

Proof (Theorem 3) Taking \( x = \hat{x} \), and using equations \( \epsilon \), \( \beta \), and \( \beta \), with \( w_k = \lambda_k \) and \( \beta_k = 0 \), and the fact that the first formula with weights \( \lambda(\hat{x}) \) interpolates \( y \) at the nodes \( \hat{w} \), we obtain

\[
\hat{f}(p(\hat{x}, \hat{y}, \hat{w})) - P(\hat{x}, \hat{y}) = \left( \prod_{k=0}^{n} (\hat{x} - \hat{x}_k) \right) \left( \sum_{k=0}^{n} \lambda_k \gamma_k \frac{(3n+5)k}{x - \hat{x}_k} \right)
\]

\[
= \left( \prod_{k=0}^{n} (\hat{x} - \hat{x}_k) \right) \left( \sum_{k=0}^{n} \lambda_k \gamma_k \frac{(3n+5)k}{x - \hat{x}_k} \right) = P(\hat{x}, \hat{y}, \hat{w}),
\]

for

\[
\lambda_k := (1 + \gamma_k) (3n+5)k - 1 = z_k (3n+5)k + (3n+5)k - 1.
\]

The definition of \( (3n+5)k \) in [10] and Lemma 1 show that

\[
1 - (3n+5) \epsilon \leq (3n+5)k \leq \frac{1}{1 - (3n+5) \epsilon},
\]

and it follows that

\[
|\lambda_k| = (3n+5) \epsilon \leq (3n+5)k \leq |\lambda_k| + (3n+5) \epsilon \leq \frac{|\lambda_k| + (3n+5) \epsilon}{1 - (3n+5) \epsilon}.
\]

and \( \epsilon \) follows from [10]. Finally, \( \epsilon \) follows from the continuity of \( P(\hat{x}, \hat{y}, \hat{w}) \), the bound on \( \nu_k \) above and \( \epsilon \).

\( \square \)

Proof (Theorem 4) Let us write \( x := \hat{x} \), consider the function

\[
g(x, y, w) := \sum_{j=0}^{n} \frac{w_j y_j}{x - \hat{x}_j},
\]

and \( \hat{w} := \lambda(\hat{x}) \). By the definition of \( z_k \) in \( \epsilon \), \( \hat{w} = \hat{w} \hat{w} \), where \( \hat{w} \hat{w} \) is the component-wise product of \( \hat{w} \) and \( \hat{z} \), i.e., \( \hat{w} \hat{w} \hat{z} = \hat{w} \hat{z} \). It follows that, for \( \nu \in \mathbb{R}^{n+1} \),

\[
g(x, \nu, \hat{w}) = \sum_{j=0}^{n} \frac{w_j y_j}{x - \hat{x}_j} = \sum_{j=0}^{n} \frac{\hat{w}_j \hat{y}_j}{x - \hat{x}_j} = \sum_{j=0}^{n} \frac{\hat{w}_j \hat{z} \hat{y}_j}{x - \hat{x}_j} = \sum_{j=0}^{n} \frac{w_j \hat{y}_j}{x - \hat{x}_j} = g(x, \hat{w}, \hat{w}) + g(x, \nu, \hat{w}),
\]

and we can write the difference \( g(x, \hat{y}, \hat{w}) - g(x, \hat{y}, \hat{w}) \) in \( \epsilon \) as

\[
S(x, y, \hat{w}) := \frac{g(x, y, \hat{w}) - g(x, y, \hat{w})}{g(x, y, \hat{w})} = \frac{g(x, y, \hat{w}) + g(x, \hat{y}, \hat{w}) - g(x, y, \hat{w}) - g(x, \hat{y}, \hat{w})}{g(x, y, \hat{w})} = \frac{N + \Delta N}{D + \Delta D}.
\]
The bounds (70), (71), (76) and Lemma 7 lead to the Error Polynomial $E(x;\hat{x},y,z)$ in (5), we get

$$S(x,y,z) = \frac{1}{D} \left( N + \frac{\Delta N}{1 + \frac{\Delta D}{D}} - N \right) = \frac{1}{1 + \frac{\Delta D}{D}} E(x;\hat{x},y,z).$$  \hspace{1cm} (92)$$

The ratio $\frac{\Delta D}{D}$ is equal to $P(x,\hat{x},z)$. Therefore, the denominator of (92) satisfies

$$1 + A_{x^2 x^2 \hat{g}} \|z\|_\infty \geq 1 + \left| \frac{\Delta D}{D} \right| \geq 1 - \frac{\Delta D}{D} \geq 1 - |P(x,\hat{x},z)| \geq 1 - A_{x^2 x^2 \hat{g}} \|z\|_\infty. \hspace{1cm} (93)$$

The forward bound (38) follows from (92) and (93) and we are done.

**Proof (Theorem 5)** Expanding the polynomials $P(x,\hat{x},y,z)$ in (5) in Lagrange's basis we obtain

$$|E(x;\hat{x},y,z)| = \left| \sum_{k=0}^n \ell_k(x)(y_k - P(x;\hat{x},y)) \right| \leq$$

$$\leq \sum_{k=0}^n |\ell_k(x)| |f(\hat{x}_k) - f(x)| + \sum_{k=0}^n |\ell_k(x)| |f(x) - P(x,\hat{x},y)| \leq$$

$$L \left( \max_{0 \leq k \leq n} |\ell_k(x)| \right) \left| x - \hat{x}_k \right| + A_{x^2 x^2 \hat{g}} \|z\|_\infty \sup_{x^2 \leq x^2 + 4} |f(x) - P(x,\hat{x},y)|.$$  \hspace{1cm} (94)

This inequality yields (17). In order to prove (48), note that, according to Theorem 16.5 in page 196 of [15] (Jackson's Theorem), there exists a polynomial $P^*$ of degree $n$ such that

$$\sup_{a \in [x^2 - 4]} |f(x) - P^*(x)| \leq \frac{L(x^2 - x^2)}{4(n+1)}. \hspace{1cm} (94)$$

Using the well known bound

$$\max_{x^2 \leq x^2 + 4} |f(x) - P(x,\hat{x},y)| \leq \left( 1 + A_{x^2 x^2 \hat{g}} \right) \sup_{x^2 \leq x^2 + 4} |f(x) - P^*(x)|$$

and (74) we deduce the that

$$A_{x^2 x^2 \hat{g}} \|z\|_\infty \sup_{x^2 \leq x^2 + 4} |f(x) - P(x,\hat{x},y)| \leq A_{x^2 x^2 \hat{g}} \left( 1 + A_{x^2 x^2 \hat{g}} \right) \|z\|_\infty \frac{L(x^2 - x^2\hat{x})}{4(n+1)},$$

and (48) follows from this last bound and (74). \hspace{1cm} \square

**Proof (Theorem 6)** The definition of Lagrange polynomial (20) leads to

$$\ell_k(x)(x-x_k) = \lambda_k(x) \prod_{k=0}^n (x-x_k).$$

In [16]'s notation, we have $\lambda_k(x) = \frac{1}{\phi_{n+1}(x)}$, and from its equations (5) and (6) we obtain $\lambda_k(x) \leq 2^{n-1}/n$. At the top of the second column in page 156 of [16] we learn that $\prod_{x=k}^n (x-x_k) \leq 2^{1-n}$ for $x \in [-1,1]$, and combining this bounds we obtain that

$$\tau(x) = \max_{x \in [-1,1]} |\ell_k(x;x^*) - x^*| \leq \left( 2^{n-1}/n \right) \times 2^{1-n} = 1/n. \hspace{1cm} (95)$$

The bounds (70), (71), (79) and (75) and Lemma 7 lead to

$$\tau(x) \|z\|_1 \leq \frac{1.0097}{n} \times 3.2704 \times 10^{-16} n^2 (2.9 + \log n)^2 \leq 3.3022 \times 10^{-16} n (2.9 + \log n)^2.$$
The effects of rounding errors in the nodes on barycentric interpolation

The bounds \( (70) \) and \((72)\) show that, in Salzer's case, for \( A = 0.67667 \) and \( B = 1.0236 \),

\[
\Lambda_{1,-q} (1 + \Lambda_{1,-q}/\hat{x}) \left\| \bar{x} \right\| \leq \left( 1 + \frac{(1 - \hat{x}^2)}{4(n+1)} \right) \leq (\log n + B)(\log n + B + 1) + 1.1328 \times 10^{-15} n^2 \pi \frac{2}{(n+1)}
\]

\[
\leq (\log n + B)(\log n + (B + 1)/A) \times 1.7794 \times 10^{-15} \leq 8.1476 \times 10^{-16}(2.9 + \log n)^2,
\]

because \((x + \hat{x}) \leq (2.9 + x)^2\) for all \( x \geq 0 \). Adding the last two bounds and using Theorem 5, we obtain

\[
|E(\hat{x}; \hat{x}, y, z)| \leq 1.1450 \times 10^{-15} \log n (2.9 + \log n)^2,
\]

and Lemma 3 and Theorem 8 lead to

\[
|q(\hat{x}; \hat{x}, \hat{y}, \hat{z}, \hat{y}) - P(\hat{x}; \hat{y}, y)| \leq 1.2042 \times 10^{-15} \log n (2.9 + \log n)^2,
\]

and we proved \((50)\). Equation \((51)\) follows from this bound, \( |\hat{x} - x| \leq 4.6 \times 10^{-6}, n \geq 10 \) and \((52)\).

**Proof (Theorem 7)** The hypothesis of this theorem asks for \( n \leq 10^6 \). Therefore, \( 2n \leq 2 \times 10^{16} \) and the bounds \((70)\) and \((72)\) for \( 2n + 1 \) nodes yield

\[
\max_{\hat{x} < x} E(x; \hat{x}, y, z) \leq hL(\log n + B + A \log 2)
\]

for \( A = 0.67667 \) and \( B = 1.0236 \). Theorem 8, \((70)\) and \((72)\) for \( n \) nodes lead to an upper bound of

\[
\frac{A}{1 - 10.841 \times 4.5312 \times 10^{-15} \left( \log n + B \right) \log 2} \leq 0.71163Lh \left( \log n + 2.2059 \right),
\]

and this leads to \((52)\). Finally, \((53)\) follows from \((99)\), \( |\hat{x} - x| \leq 4.6 \times 10^{-6} \) and Theorem 7.

**Proof (Theorem 6)** The hypothesis \( \| \lambda(\lambda, \hat{w}) \|_\infty (1 + \Lambda_{1,-q}/\hat{x}) \leq 1 \) allows us to use theorem 3 in \((13)\) with \( d = 0, x = \hat{x} \) and \( w = \lambda(\hat{x}) \) and conclude that the Lebesgue constant \( A_{1,-q, \hat{w}} \), which is defined in the more general context of rational interpolation in \((13)\), satisfies \( A_{1,-q, \hat{w}} \leq 1 \). The hypothesis \((n+2)(2+E) \leq 1 \) and Theorem 1 in \((13)\) with \( x = \hat{x} \) and \( w = \hat{w} \) lead to

\[
\|q(\hat{x}; \hat{x}, \hat{y}, \hat{w}) - q(\hat{x}, \hat{x}, \hat{y}, \hat{w})\|_\infty \leq hL(\log n + B + A \log 2)
\]

with

\[
\hat{y}_k = y_{k+1}(1 + \alpha_k)(1 + v_k),
\]

where

\[
\| \theta \|_\infty \leq \frac{\theta(2n+5)}{1 - (2n+5)\theta} \| v \|_\infty \epsilon,
\]

and \( \| y - y \|_\infty \leq 1476 \times 10^{-15} \log n \). In Salzer's case, Lemmas 3 and 4 show that

\[
A \leq 1 - 4.5312 \times 10^{-15} \times 1.7794 \leq 0.71504 \times (1.5128 + \log n) \leq 11.456,
\]

and noting that \( n + 2 \leq (2n + 5)/2 \) and using this bound on \( A \) we deduce that

\[
\| q(\hat{x}; \hat{x}, \hat{y}, \hat{w}) - q(\hat{x}, \hat{x}, \hat{y}, \hat{w}) \|_\infty \leq 0.35754 \times \log n (2n + 5) \| y - y \|_\infty \epsilon,
\]

and \((55)\) follows from the last two inequalities.

**Proof (Theorem 7)** We start by showing that for each \( 0 \leq k \leq n \) there exists \( \epsilon_k \) such that

\[
\epsilon_k + 1 = (x_{k+1} + 1)(1 + \epsilon_k) \quad \text{and} \quad |\epsilon_k| \leq \| \theta \|_\infty.
\]

We analyze the \( x_k \) in the three bins:
If $x_k$ is in the first bin then $\epsilon_k = \theta_k$ is appropriate, because $x_k + 1 = (1 + \theta_k) r_k = (x_k + 1)(1 + \theta_k)$.

If $x_k$ is in the second bin then $2|x_k| \leq 1$ and

$$
x_k + 1 = 1 + x_k (1 + \theta_k) = (x_k + 1)(1 + \epsilon_k)
$$

for $\epsilon_k := \frac{x_k \theta_k}{1 + x_k}$.

This $\epsilon_k$ satisfies (60) because $|x_k/(1 + x_k)| \leq 1$ when $2|x_k| \leq 1$.

If $x_k$ is in the third bin then $\epsilon_k = 1 + \epsilon_k = 1 + r_k (1 + \theta_k)$, with $r_k = x_k - 1$. Thus,

$$
x_k + 1 = 2 + (x_k - 1)(1 + \theta_k) = (x_k + 1)(1 + \epsilon_k)
$$

for $\epsilon_k := \frac{x_k - 1}{x_k + 1}$.

This $\epsilon_k$ is valid because $0 \leq x_k \leq 1$ implies that $-\frac{4}{x_k + 1} \leq 0$.

By symmetry, we need to verify (60) only for $0 \leq k \leq n/2$. Let us then assume from now on that $0 \leq k \leq n/2$ and $0 \leq j \neq k \leq n$. Defining

$$
\delta'_k := \frac{x_j - x_k}{x_j - x_k} - 1,
$$

we obtain

$$
|\delta'_k| = \left| \frac{1}{|x_j - x_k|} \left[ (x_k + 1) - (x_k + 1) - (x_j + 1) - (x_j + 1) \right] \right|

= \left| \frac{x_j - x_k}{x_j - x_k} \right| \leq \theta \left( 1 + 2 \frac{x_k + 1}{|x_j - x_k|} \right).
$$

Since we are assuming that $-1 \leq x_k \leq 0$, it follows from (60) and (10) that

$$
|\delta'_k| = \left| \frac{\delta'_k}{1 + \delta'_k} \right| \leq \psi \left( 1 + 2 \frac{x_k + 1}{|x_k - x_j|} \right)
$$

for

$$
\psi := \frac{\|\theta\|_\infty}{1 - \max_{0 \leq j, k \leq n} |\delta'_k|} \leq \frac{\|\theta\|_\infty}{1 - \|\theta\|_\infty (1 + 2 \times 0.20432 \times n^2)} \leq 1.0008 \|\theta\|_\infty.
$$

In particular, if $k = 0$ then $x_k = -1$ and this equation shows that $|\delta'_k| \leq \psi$ for all $j$. This implies that

$$
\sum_{j \neq k} |\delta'_k| \leq \eta \psi \text{ and (50) is satisfied for } k = 0.
$$

Therefore, from now on we assume that $k > 0$.

Combining equation (99) with Lemma (10) we obtain

$$
\frac{1}{\psi} \sum_{j \neq k} |\delta'_k| \leq n + 2 (x_k + 1) \left( 1 + k \log(4k - 1) \right) \frac{1 + 4n}{2n} \log(4k + 1) + 3 \left( \frac{1}{2} + \frac{4}{2} \right).
$$

Let us now analyze the case $1 \leq k \leq \frac{n}{4}$. The definition $x_k := -\cos \frac{k \pi}{2n}$ yields

$$
x_k + 1 = 2 \sin^2 \left( \frac{k \pi}{2n} \right).
$$

Equation (101), the assumption $k \leq n/3$ and the inequality $\sin \left( \frac{4 \pi}{3} \right) \geq \sin \left( \frac{4 \pi}{3} \right)$ yield

$$
\frac{1}{\psi} \sum_{j \neq k} |\delta'_k| \leq n + 2 (1 + k \log(4k - 1)) + 3 \pi k \frac{1}{4} \log(4k + 1)
$$

$$
\leq n + 2 + n \left( \frac{2}{3} + \pi^2 \right) \frac{1}{4 \sqrt{2}} \left( \frac{4n}{2} + n \right) \log \left( \frac{4n}{2} + \frac{3}{2} \right) \leq 2 + 1.8n + 2.5 n \log(n + 1)
$$

and (50) holds for $1 \leq k \leq n/3$. Finally, if $n/3 < k \leq n/2$, then equation (99) and Lemma (10) lead to

$$
\frac{1}{\psi} \sum_{j \neq k} |\delta'_k| \leq n + \frac{n^2}{2 \sqrt{2}} \left( 1 + \log(4k - 1) \right) + 3 \log(4k + 1) + \frac{1}{k^2} \left( \frac{1}{k^2} + \log(4k - 1) \right).
$$
The expression above decreases as $k$ increases for $k \geq 1/3$. Therefore, replacing $k$ by $n/3$ we get

$$\frac{1}{\psi} \sum_{j=1}^{n} |\delta_{jk}| \leq n + \frac{n^2}{2\sqrt{2}} \left( \frac{9}{n^2} + \frac{3}{\sqrt{2}} \log \left( \frac{2n}{3} + 1 \right) + \frac{9}{\sqrt{2}} \log \left( \frac{4n}{3} + 1 \right) \right)$$

$$= n + \frac{9}{2\sqrt{2}} \log \left( \frac{4n}{3} + 1 \right) \leq 3.1820 + n + 4.2427n \log \left( \frac{4n}{3} + 4 \right)$$

$$\leq 3.1820 + 2.2213n + 4.2427n \log(n + 1).$$

This equation and (100) lead to (56) and we are done.

## B Experimental details

The data in our tables and plots were generated with C++ code compiled in Ubuntu 12.04 with gcc 4.8.1 with options -m64 -std=c++11 -Ofast -Wall, with NDEBUG defined. The code was executed on an Intel Core i7 2700K processor. We used the IEEE-754 double precision arithmetic, with C++'s type double. These numbers are used by Matlab and correspond to real*8 in Fortran.

We used gcc 4.8.1's quadruple precision type __float128 as a benchmark: we considered results obtained using this arithmetic as exact. We checked our results with Intel's C++ 13.0 compiler with option -Qoption,cpp,-extended_float_type and its quadruple precision type __Quad. We also performed accuracy experiments with the versions of the compilers above for Windows 7 and OS X, in a Quad Core Intel Xeon processor. There were no relevant differences in the results.

The sets $X_{-1,0}$ and $X_{0,1}$ in tables 1 and 2 contain $10^3$ points each. These points are distributed in 100 intervals $[x_{1}, x_{20}]$. $X_{-1,0}$ uses $0 < x < 100$ and $X_{0,1}$ considers $n/2 - 100 < x < n/2$. In each interval $(x_{k-1}, x_{k})$ we picked the 200 floating point numbers to the right of $x_{k}$ and the 200 floating point numbers to the left of $x_{k-1}$. The remaining 600 points where equally spaced in $(x_{k-1}, x_{k+1})$. The Step II errors in tables 1 and 2 were estimated by performing Step II in double precision, Step III was evaluated with gcc 4.8.1's __float128 arithmetic, and the result was compared with the interpolated function evaluated with __float128 arithmetic.

The trial points in Figure 2 were chosen as follows: for each $n$ we considered the relative errors $y_{k}$ and $y_{k}^{*}$ corresponding to the Salzer’s weights and the numerical weights, we then picked 4 groups of ten indexes: the ten indexes corresponding largest values $y_{k}$, the ten indexes corresponding to the ten smallest values of $y_{k}^{*}$ and the analogous 20 indexes for $y_{k}$. We then removed the repeated indexes and obtained a vector with $n_{i}$ indexes. For each index $k > 0$ we picked the 2000 floating points to the left of $x_{k}$ and for each index $k < n$ we picked the 2000 floating points to the right of $x_{k}$. We then considered the $n_{i}$ intervals of the form $[x_{k-1}, x_{k}]$ for $k = 1, n/n_{i}, 2n/n_{i}, \ldots$ and picked 2000 equally spaced points in each of these intervals. The first formula was evaluated at these trial points using the procedure suggested in subsection 3.1 our article [12].

The $b_{0}$ in Table 3 were computed using __float128 arithmetic, from nodes $x_{k}$ obtained from the formula $x_{k} = \sin \left( \frac{k}{n} \pi \right)$ using IEEE-754 double arithmetic. The versions with 3 bins in tables 4 and 5 are as in Figure 6. The versions with 39 bins consider the central bin $[-2^{-30}, 2^{-30}]$, with base $b_{39} = 0$, the bins $[-1, 2^{-30}]$ for $2 \leq k < 10$ and $[-2^{-8}, -2^{-k-1}]$ for $1 \leq k < 9$, with base at their left extreme point. The remaining 19 bins and bases were obtained by reflection around 0. The versions with 79 bins consider the central bin $[-2^{-30}, 2^{-30}]$, with base $b_{79} = 0$, the bins $[-1, 2^{-20}]$, $[-2^{-k}, 2^{-k+1}]$ for $1 \leq k < 20$ and $[-2^{-8}, -2^{-k-1}]$ for $1 \leq k < 19$, with base at their left extreme point. The remaining 39 bins and bases were obtained by reflection.

The times in Table 3 were measured using the cpu timer available in version 1.54.0 of the boost library 4. This timer measures the time taken only by the process one is concerned with.

## References

1. J. –P. Berrut, 1988, Rational functions for guaranteed and experimentally well conditioned global interpolation, Comput. Math. Appl. 15 pp 1-16.
2. L. Bos, S. De Marchi and K. Hormann, 2011, On the Lebesgue constant of Berrut’s rational interpolant at equidistant nodes, J. Comput. Appl. Math., 236(4), 504–510.
3. L. Bos, S. De Marchi, K. Hormann and J. Sidon, 2013, Bounding the Lebesgue Constant for Berrut’s rational interpolant in general nodes. J. Approx. Theory, 169, 7-22.
4. The boost library website: [http://www.boost.org/]
5. T. J. Dekker, 1971, A floating-point technique for extending the available precision, Numer. Math., 18, 224–242.
6. M. Floater and K. Hormann, 2007, Barycentric rational interpolation with no poles and high rates of approximation, Numer. Math. 107, 315-331.
7. D. Goldberg, 1991, What Every Computer Scientist Should Know About Floating Point Arithmetic, ACM Comput. Surv., 23, n. 1, 5–48.
8. R. Giannotti, 1980, Evaluation of Lebesgue constants, SIAM J. Numer. Anal. 17 (4) pp. 512–520.
9. N. J. Higham, 2004, The numerical stability of barycentric Lagrange interpolation, IMA J. Numer. Anal. 24, 547–556.
10. N. Higham, 2002, Accuracy and Stability of Numerical Algorithms, 2nd ed., SIAM, Philadelphia.
11. IEEE Computer Society, 2008, IEEE Standard 754-2008 for floating-point arithmetic.
12. W. F. Mascarenhas, 2014, The stability of barycentric interpolation at the Chebyshev points of the second kind, Numer. Math., available online, DOI: 10.1007/s00211-014-0612-6.
13. W. F. Mascarenhas and A. Pierro de Camargo, On the backward stability of the second barycentric formula for interpolation, submitted to publication.
14. J. McCabe and G. Phillips, 1973, On a certain class of Lebesgue constants, BIT Numer. Math., 13 (4) 434–442.
15. M. J. D. Powell, 1981, Approximation theory and methods, Cambridge University Press.
16. H. Salzer, 1972, Lagrangian interpolation at the Chebyshev points $x_n \equiv \cos(\nu \pi/n)$, $\nu = 0(1)n$; some unnoted advantages, Comput. J., 15(2), pp. 156–159.
17. W. J. Taylor, 1945, Method of Lagrangian curvilinear interpolation, J. Res. Nat. Bur. Standards, v. 35, pp 151-155.
18. W. Werner, 1984, Polynomial interpolation: Lagrange versus Newton, Math. Comp. 43, pp 205-207.