Error Analysis of Finite Differences and the Mapping Parameter in Spectral Differentiation

Divakar Viswanath

Department of Mathematics, University of Michigan (divakar@umich.edu).

Abstract

The Chebyshev points are commonly used for spectral differentiation in non-periodic domains. The rounding error in the Chebyshev approximation to the \( n \)-the derivative increases at a rate greater than \( n^{2m} \) for the \( m \)-th derivative. The mapping technique of Kosloff and Tal-Ezer (J. Comp. Physics, vol. 104 (1993), p. 457-469) ameliorates this increase in rounding error. We show that the argument used to justify the choice of the mapping parameter is substantially incomplete. We analyze rounding error as well as discretization error and give a more complete argument for the choice of the mapping parameter. If the discrete cosine transform is used to compute derivatives, we show that a different choice of the mapping parameter yields greater accuracy.

1 Introduction

The Chebyshev points \( x_j = \cos(j\pi/n) \), \( n = 0, 1, \ldots, n \), are commonly used to discretize the interval \([-1, 1]\) when the boundary conditions are not periodic. The \( m \)-th derivative \( f^{(m)}(x) \) may be approximated as \( \sum_{k=0}^{m} f(x_k)w_{k,m} \) where \( w_{k,m} \) are differentiation weights. The rounding error in the \( m \)-th derivative increases faster than \( n^{2m} \) (precise asymptotics will be given in section 2). In contrast, the rounding error error in Fourier spectral methods increases at the much milder rate of \( n^m \) or \( n^{m+1} \).

Kosloff and Tal-Ezer [7] introduced a mapping technique to control the growth in rounding errors while preserving spectral accuracy. The central idea is to replace the function \( f(x) \) by the function \( F(\xi) = f(g(\xi)) \) where \( g : [-1, 1] \to [-1, 1] \), where

\[
g(\xi) = \frac{\arcsin \alpha \xi}{\arcsin \alpha}
\]

is a mapping function that depends upon the parameter \( \alpha \in [0, 1] \). The grid in \( \xi \) is still Chebyshev with \( \xi_j = \cos(j\pi/n) \), and is used to define the mapped grid in \( x \) as \( x_j = \xi_j \) for \( j = 0, 1, \ldots n \). The derivative is approximated using

\[
\frac{df}{dx} = \frac{1}{g'(\xi)} \frac{dF}{d\xi}.
\]
The derivative $dF/d\xi$ is obtained using spectral differentiation at Chebyshev points and then scaled by $1/g'(\xi)$ to obtain $df/dx$. Higher derivatives are obtained by iteration of this technique.

The points $x_j$ converge to Chebyshev and equi-spaced points, respectively, in the limits $\alpha \to 0$ and $\alpha \to 1$. For $\alpha$ in-between, and usually quite close to 1, the grid is nearly equi-spaced and still retains spectral accuracy. Since the grid points are not clustered quadratically near the endpoints $\pm 1$, the growth of rounding errors is milder [1] [7].

The function $F(\xi)$ will have a singularity in the complex plane, due to the mapping, even if $f(x)$ is an entire function. Inspection of (1.1) shows that there are singularities at $\xi = \pm 1/\alpha$. If $f(x)$ is an entire function, such as $f(x) = \sin Kx$, the interpolation error in $F(\xi)$ using Chebyshev points and in $f(x)$ using the mapped grid are both controlled by the singularity locations $\pm 1/\alpha$. Kosloff and Tal-Ezer [7] recommended the choice of $\alpha$ determined by

$$\left( \frac{1 - \sqrt{1 - \alpha^2}}{\alpha} \right)^n = u$$

(1.2)

where $u$ is the desired accuracy. Don and Solomonoff [4] showed that taking $u$ to be the machine precision leads to accurate derivatives. We prefer to take $u$ to be the unit roundoff (for double precision arithmetic, the unit roundoff is $u = 2^{-53}$ and the machine epsilon is $2^{-52}$ [5]) because $u$ is the quantity that comes up naturally in rounding error analysis. However, the distinction between unit roundoff and machine epsilon has no real consequence in this situation. The solution of (1.2) is given by $\alpha = 2/(t + 1/t)$ with $t = u^{-1/n}$.

A plausible argument for (1.2) is that it balances the discretization error on the left hand side with the rounding error on the right hand side. Balancing errors is the right idea, but it begs the question of why the $n^{2m}$ or $n^{2m+1}$ increase in rounding error is not showing up in (1.2). In this article we give a systematic treatment of both rounding and discretization errors and show that (1.2) is still the right equation regardless of the order of the differentiation $m$. The order of differentiation $m$ introduces prefactors into both discretization and rounding error, and these cancel off fortuitously to leave (1.2) as the correct equation for the mapping parameter $\alpha$ regardless of $m$.

Computation of derivatives at Chebyshev points incurs more error when the discrete cosine transform is used [4], in comparison with carefully computed differentiation matrices [1] [4]. However, the discrete cosine transform is much faster. We show that (1.2) can be modified to choose $\alpha$ in a way that yields slightly more accurate derivatives when the discrete cosine transform is employed.

Sections 2 and 3 present analyses of rounding and discretization errors, respectively, showing how the pre-factors cancel leading to (1.2). When $n$ is small the total error is dominated by discretization error and when $n$ is large the total error is dominated by rounding error. In section 3, we show that the value of $n$ at which the total error transitions from discretization error to rounding error does not depend upon $m$, the order of differentiation.

In section 4, we specialize arguments to the mapping (1.1). We consider the slightly more general balancing equation

$$\left( \frac{1 - \sqrt{1 - \alpha^2}}{\alpha} \right)^n = n^\beta u$$

(1.3)

and find that $\beta = 0$ is a good choice when accurate differentiation matrices are used and $\beta = 0.5$ is a better choice for the discrete cosine transform.
2 Rounding error analysis of finite differencing

Spectral differentiation at Chebyshev points is a special case of finite differencing. In this section, we derive rounding error bounds assuming the method of partial products. The method of partial products is an efficient way to calculate finite difference weights \[9\]. The rounding error bounds here include the errors that arise during the calculation of finite difference weights. Some quantities that arise will recur in the analysis of discretization error. Comparison to rounding error bounds which assume that the finite difference weights are exact shows that computation of finite difference weights introduces only a modest amount of error. Finally we give asymptotic estimates of the error in the limit \(n \to \infty\).

For floating point arithmetic, we mostly follow Higham \[5\], with a few modifications from \[8\]. The axiom of floating point arithmetic is \(fl(x, y) = (x, y)(1 + \delta)\) with \(|\delta| \leq u\), where \(u\) is the unit-roundoff (\(2^{-53}\) for double precision arithmetic). To handle the accumulation of rounding error, we denote \((1 + \delta_1)^{\rho_1}(1 + \delta_2)^{\rho_2}\ldots(1 + \delta_n)^{\rho_n}\), with each \(\rho_i\) equal to \(+1, 0, -1\) and \(|\delta_i| \leq u\), by \(1 + \theta_n\). In our convention, each occurrence of \(\theta_n\) is local, which means that two occurrences of \(\theta_n\), even in the same equation, are not assumed to be equal. The quantity \(\theta_n\) stands for any quantity that may be realized as the accumulated relative error of \(n\) or fewer multiplications and divisions. It satisfies \(|\theta_n| \leq \gamma_n\), where \(\gamma_n = nu/(1-nu)\), as long as \(nu < 1\). Whenever \(\gamma_n\) occurs, it is implicitly assumed that \(nu < 1\).

Computed quantities are hatted. Thus if \(s = x_1 + \cdots + x_n\), with each \(x_i\) a floating point number, the computed quantity is denoted \(\hat{s}\). If the addition is from left to right, we may write

\[
\hat{s} = x_1(1 + \theta_{n-1}) + x_2(1 + \theta_{n-1}) + \cdots + x_n(1 + \theta_1).
\]

Conventions stated above allow us to rewrite this as

\[
\hat{s} = x_1(1 + \theta_{n-1}) + x_2(1 + \theta_{n-1}) + \cdots + x_n(1 + \theta_{n-1}).
\]

This device will be employed frequently. Notice that it is a mistake to factor out \((1 + \theta_{n-1})\) in the right hand side, because each \(\theta_{n-1}\) is a local variable and two distinct instances are not necessarily equal. However, we may write \(\hat{s}\) as \(\sum x_j(1 + \theta_{n-1})\), with the assumption that each \(\theta_{n-1}\) inside the summation is different.

2.1 Bounds for rounding error

Assume the \(n + 1\) grid points to be \(x_0, x_1, \ldots, x_n\). The weight \(w_{k,m}\) in the finite difference formula \(f^{(m)}(x) = \sum_{k=0}^{n} w_{k,m} f(x_k) + \text{error}\) is given by

\[
w_{k,m} = \frac{d^m \ell_k(x)}{dx^m} = w_k \frac{d^m}{dx^m} \prod_{j=0, j \neq k}^{n} (x - x_j),
\]  \hspace{1cm} (2.1)

where \(\ell_k(x)\) is the Lagrange cardinal function \(\prod_{j \neq k} (x - x_j) / \prod_{j \neq k} (x_k - x_j)\) and \(w_k\) is the Lagrange weight \(1/\prod_{j \neq k} (x_k - x_j)\).

If we assume \(x = 0\), by shifting the grid if necessary, then

\[
w_{k,m} = (-1)^{n-m} m! w_k S_{n-m} \{x_0, \ldots, x_n\} - \{x_k\},
\]  \hspace{1cm} (2.2)
where $S_{n-m}$ is the elementary symmetric function of order $n-m$ [9]. The elementary symmetric function $S_{n-m}$ is the sum of $\binom{n}{n-m}$ terms each of which is a product of a selection of $n-m$ entries out of the $n$ (all grid points excluding $x_k$). $S_0$ is defined as 1.

In the method of partial products [9], the weight $w_{k,m}$ is computed as follows. The polynomials $\prod_{j=0}^k (x-x_j)$ and $\prod_{j=k}^n (x-x_j)$ are denoted by $L_k$ and $R_k$, respectively. Define

$$w_{k,m} = \text{coeff of } x^m \text{ in } L_{k-1}R_{k+1} = (-1)^{n-m} \sum_{m_1m_2} S_{k-m_1} (x_0, \ldots, x_{k-1}) S_{n-k-m_2} (x_{k+1}, \ldots, x_n), \quad (2.3)$$

where the sum is taken over nonzero integers $m_1, m_2$ satisfying $m_1 + m_2 = m$, $k-m_1 \geq 0$, and $n-k-m_2 \geq 0$. The finite difference weight $w_{k,m}$ is obtained as $m!w_kw_{k,m}'$, where $w_k$ is the Lagrange weight at $z_k$.

The elementary symmetric functions that appear in (2.3) are computed by forming the products $L_k$ and $R_k$, recursively [9]. In effect the recurrence

$$S_{N-m}(y_1, \ldots, y_N) = \begin{cases} y_N S_{N-1}(y_1, \ldots, y_{N-1}) & \text{if } m = 0 \\ S_{N-m}(y_1, \ldots, y_{N-1}) + y_N S_{N-m-1}(y_1, \ldots, y_{N-1}) & \text{if } N > m > 0 \\ 1 & \text{if } m = N \end{cases} \quad (2.4)$$

is used for the computation of symmetric functions.

To prove an upper bound on the rounding error in computing $\sum_{k=0}^n w_{k,m}f(x_k)$, we begin with the following lemma.

**Lemma 1.** If the recurrence (2.4) is used to calculate $S_{N-m}(y_1, y_2, \ldots, y_N)$, the computed quantity may be represented as

$$\hat{S}_{N-m} = \sum_{i_1 < \cdots < i_{N-m}} y_{i_1}y_{i_2}\cdots y_{i_{N-m}}(1 + \theta_{f(N,m)})$$

with $f(N,m) = 2(N-1) - m$ for $0 \leq m \leq N$.

**Proof.** One may easily verify that $f(1,0) = f(2,0) = f(1,1) = 0$ and $f(2,0) = f(2,1) = 1$ suffice. If we inductively assume the lemma for $S_{N-m}(y_1, \ldots, y_{N-1})$ and $S_{N-m-1}(y_1, \ldots, y_{N-1})$, and apply the floating point axiom to the recurrence, we get

$$f(N,m) \leq \max (f(N-1,m-1) + 1, f(N-1,m) + 2)$$

for $N > 2$, along with $f(N,N) = 0$ and $f(N,0) = 1 + f(N-1,0)$. It may be easily verified that $f(N,m) = 2(N-1) - m$ satisfies these relations.

Next we turn to the roundoff analysis of $w_{k,m}'$ computed using (2.3).

**Lemma 2.** The computed value of $w_{k,m}'$ may be represented as

$$\hat{w}_{k,m}' = (-1)^{n-m} \sum_{i_1 < \cdots < i_{n-m}} x_{i_1}x_{i_2}\cdots x_{i_{n-m}}(1 + \theta_{2n+1})$$

where the summation is over $i_j \in \{0,1,\ldots,n\} - \{k\}$.
Proof. The number of terms in the summation in (2.3) is at most \(m + 1\) and each term is formed using a single multiplication. Therefore we may represent the computed quantity as

\[
(-1)^{n-m} \sum_{m_1, m_2} \hat{S}_{k-m_1}(x_0, \ldots, x_{k-1}) \hat{S}_{n-k-m_2}(x_{k+1}, \ldots, x_n)(1 + \theta_{m+1}).
\]

Applying Lemma (1) to \(\hat{S}_{k-m}\) (with \(N = k\) and \(m = m_1\)) and \(\hat{S}_{n-k-m_2}\) (with \(N = n - k\) and \(m = m_2\)), we get a representation of \(\hat{w}_{k,m}'\) that completes the proof.

The following lemma occurs as a part of Higham’s rounding error analysis of the barycentric formula [6].

Lemma 3. The computed Lagrange weight \(\hat{w}_k\) is given by

\[
\hat{w}_k = w_k(1 + \theta_{2n})
\]

where \(w_k\) is the exact Lagrange weight.

Proof. The exact Lagrange weight is given by

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

The \(\theta_{2n}\) in the lemma is a result of \(n\) subtractions, \(n - 1\) multiplications, and a single division.

Lemma 4. The computed weight \(w_{k,m}\) may be represented as

\[
\hat{w}_{k,m} = (-1)^{n-m}m!w_k \sum_{i_1 < \cdots < i_{n-m}} x_{i_1} x_{i_2} \ldots x_{i_{n-m}}(1 + \theta_{4n+3})
\]

where the summation is over \(i_j \in \{0, 1, \ldots, n\} - \{k\}\).

Proof. The finite-difference weight \(w_{k,m}\) is computed as \(m!w_k w_{k,m}\). This lemma is proved using the previous two lemmas and incrementing the subscript of \(\theta\) by 2 to account for multiplication by \(m!\) and \(w_k\).

Lemma 5. If the derivative is being approximated at \(x = \zeta\), the computed weight \(w_{k,m}\) may be represented as

\[
\hat{w}_{k,m} = (-1)^{n-m}m!w_k \sum_{i_1 < \cdots < i_{n-m}} (x_{i_1} - \zeta)(x_{i_2} - \zeta)\ldots(x_{i_{n-m}} - \zeta)(1 + \theta_{5n-m+3})
\]

where the summation is over \(i_j \in \{0, 1, \ldots, n\} - \{k\}\).

Proof. The finite difference weights are computed at \(x = \zeta\) by shifting the grid by \(-\zeta\) and then using the algorithm for \(x = 0\). Thus compared to the previous lemma, the subscript of \(\theta\) is incremented by \(n - m\) to allow for \(n - m\) subtractions inside the summation. There is no need to redo the analysis of \(w_k\) because \(w_k\) is unchanged by the shift and it is assumed that \(w_k\) is computed prior to shifting.
Theorem 6. The magnitude of the roundoff error in the computation of the finite difference approximation

\[ \sum_{k=0}^{n} w_{k,m} f(x_k) \]

to \( f^{(m)}(\zeta) \) is upper bounded by

\[ U_R = \gamma_{6n-m+4} |f| \sum_{k=0}^{n} m! |w_k| S_{n-m} (\{|x_0 - \zeta|, \ldots, |x_n - \zeta|\} - \{|x_k - \zeta|\}) \],

(2.5)

where \( |f| \) is equal to \( \max_j |f(x_j)| \).

Proof. For the computed value of \( w_{k,m} \), we may use the previous lemma. In forming the sum \( \sum_{k=0}^{n} w_{k,m} f(x_k) \), a total of \( n+1 \) terms are added and each term is formed through a single multiplication. Therefore the computed value of \( \sum_{k=0}^{n} w_{k,m} f(x_k) \) is

\[ \sum_{k=0}^{m} f(z_k)(-1)^{n-m} m! w_k \sum_{i_1 < \cdots < i_{n-m}} (x_{i_1} - \zeta)(x_{i_2} - \zeta) \cdots (x_{i_{n-m}} - \zeta) (1 + \theta_{6n-m+4}). \]

Here \( (1 + \theta_{6n-m+4}) \) is obtained from \( (1 + \theta_{5n-m+3})(1 + \theta_{n+1}) \). The upper bound is obtained by subtracting the true value of \( \sum_k w_k f(x_k) \), taking absolute values, and using \( |\theta_{6n-m+4}| \leq \gamma_{6n-m+4} \).

If the weights \( w_{k,m} \) are exact, except for the inevitable roundoff in floating point representation, the computed value of \( \sum_{k=0}^{n} w_{k,m} f(x_k) \) is

\[ \sum_{k=0}^{n} w_{k,m} f(x_k)(1 + \theta_{k+2}) \]

assuming right to left summation. Thus the magnitude of the rounding error is bounded by

\[ U'_R = |f| \sum_{k=0}^{n} |w_{k,m}| \gamma_{k+2}. \]

(2.6)

For other orders of summation the \( \gamma_{k+2} \) may be reordered.

2.2 Asymptotics

To obtain asymptotics for \( U_R \) and \( U'_R \) in the limit of increasing \( n \), we introduce three quantities \( \mathcal{W}_\ell, \mathcal{E}_m^\ell, \) and \( \mathcal{E}_m^{\ell,k} \). The first of these \( \mathcal{W}_\ell \) is defined as \( \prod_{j=0,j\neq \ell}^{n} (x_\ell - x_j) \). It is the inverse of the Lagrange weight. If \( x_0, x_1, \ldots, x_n \) are the Chebyshev points, it is well-known (see [4] for example) that

\[ \mathcal{W}_\ell = \left\{ \begin{array}{ll} (-1)^{\ell} \frac{2n}{2n-\ell} & \text{for } \ell = 0, n \\ (-1)^{\ell} \frac{n}{2n-\ell} & \text{otherwise.} \end{array} \right. \]

(2.7)

Define

\[ \mathcal{E}_m^\ell = \sum_{i_1 < \cdots < i_m} \frac{1}{(x_\ell - x_{i_1})(x_\ell - x_{i_2}) \cdots (x_\ell - x_{i_m})} \]

(2.8)
where \( i_j \in \{0, 1, \ldots, n\} - \{\ell\} \). If \( \ell \neq k \), define

\[
\xi_m^{\ell,k} = \sum_{i_1 < \cdots < i_m} \frac{1}{(x_\ell - x_{i_1})(x_\ell - x_{i_2})\cdots(x_\ell - x_{i_m})}
\]  

(2.9)

where \( i_j \in \{0, 1, \ldots, n\} - \{k, \ell\} \). If \( m = 0 \), both \( \xi_m^\ell \) and \( \xi_m^{\ell,k} \) are defined to be 1.

Define

\[
P_r^\ell = \sum_{j=0, j \neq \ell}^n \frac{1}{(x_\ell - x_j)^r} \quad \text{and} \quad P_r^{\ell,k} = P_r^\ell - \frac{1}{(x_\ell - x_k)^r}.
\]

For the Chebyshev points, or indeed for any set of points, the rounding errors are maximized at the edges as evident from inspection of (2.5). Later we will see that discretization errors too tend to be the greatest at the edges. Therefore we set \( \ell = 0 \), and find that

\[
\frac{1}{x_\ell - x_j} = \frac{1}{1 - x_j} \sim \frac{2n^2}{j^2\pi^2}.
\]

It follows that

\[
P_r^0 \sim \frac{2^r \zeta(2r)}{\pi^{2r}} n^{2r},
\]

(2.10)

and

\[
P_r^{0,k} \sim \frac{2^r}{\pi^{2r}} n^{2r} \left( \zeta(2r) - \frac{1}{k^{2r}} \right),
\]

(2.11)

where \( \zeta(\cdot) \) is the zeta function.

The Newton identities relating symmetric functions give

\[
\xi_1^0 = P_1^0
\]

\[
\xi_0^0 = \xi_1^0 P_1^0 - P_2^0
\]

\[
\xi_3^0 = \xi_2^0 P_1^0 - \xi_1^0 P_2^0 + P_3^0.
\]

(2.12)

Similar identities related \( \xi_r^{0,k} \) and \( P_r^{0,k} \).

**Theorem 7.** If \( x_0, x_1, \ldots, x_n \) are the Chebyshev points, the upper bound \( U_R \) for the rounding error with \( \zeta = x_0 = 1 \) has the following asymptotics in the limit of increasing \( n \):

\[
U_R \sim \gamma_{6n+3}|f| \left( \frac{n^2}{3} + \sum_{k=1}^n \frac{4n^2}{\pi^2} \right) = \gamma_{6n+3}|f|0.9995 \ldots n^2
\]

\[
\sim \gamma_{6n+2}|f| \left( \frac{n^4}{30} + \sum_{k=1}^n \frac{4n^4}{\pi^4 k^2} \left( \frac{2}{3} - \frac{2}{\pi^2 k^2} \right) \right) = \gamma_{6n+2}|f|0.1665 \ldots n^4
\]

\[
\sim \gamma_{6n+1}|f| \left( \frac{n^6}{630} + \sum_{k=1}^n \frac{2n^6}{15\pi^6 k^6} |\pi^4 k^4 - 20\pi^2 k^2 + 120| \right) = \gamma_{6n+1}|f|0.01109 \ldots n^6
\]

\[
\sim \gamma_{6n}|f| \left( \frac{n^8}{22680} + \sum_{k=1}^n \frac{2n^8}{315\pi^8 k^8} |\pi^6 k^6 - 42\pi^4 k^4 + 840\pi^2 k^2 - 5040| \right) = \gamma_{6n}|f|0.00039 \ldots n^8,
\]

for order of differentiation \( m = 1, 2, 3, 4 \), respectively. As before, \( |f| = \max_j f(x_j) \).

7
Proof. If we go back to (2.5), which defines $U_R$, and look at the $k = 0$ term with $\zeta = x_0 = 1$, it can be written as
\[
w_0 S_{n-m} (1 - x_1, 1 - x_2, \ldots, 1 - x_n) = \frac{S_{n-m} (1 - x_1, 1 - x_2, \ldots, 1 - x_n)}{\prod_{j=1}^{n} (1 - x_j)} = \epsilon_0^0.
\]
A term with $k > 0$ may be written as
\[
|w_k| S_{n-m} \left( \{0, 1 - x_1, \ldots, 1 - x_n\} - \{1 - x_k\} \right) = \frac{S_{n-m} \left( \{1 - x_1, \ldots, 1 - x_n\} - \{1 - x_k\} \right)}{\prod_{j=0, j\neq k}^{n} |x_k - x_j|} = \frac{|W_0| S_{n-m} \left( \{1 - x_1, \ldots, 1 - x_n\} - \{1 - x_k\} \right)}{|W_k| \prod_{j=1}^{n} (1 - x_j)} = \frac{|W_0|}{|W_k|} \epsilon_{m-1}^{0,k}.
\]
So the summation in (2.5) becomes
\[
\epsilon_m^0 + \sum_{k=1}^{n} \frac{|W_0|}{|W_k|} \epsilon_{m-1}^{0,k}.
\]
(2.13)

The proof is completed using the asymptotics for $P_r^0$ and $P_r^{0,k}$ in (2.10) and (2.11), along with Newton identities (2.12), to obtain the asymptotics of $\epsilon_m^0$ and $\epsilon_{m-1}^{0,k}$. These along with the formula (2.7) for $W_\ell$ are substituted into (2.13) to obtain the asymptotics of that quantity.

The methods that utilize accurate versions of the spectral differentiation matrix [1, 4] are often employed with $m = 1$. Therefore we limit the next theorem to $m = 1$.

**Theorem 8.** If $m = 1$ and $\zeta = x_0 = 1$, the upper bound $U'_R$ defined by (2.6) satisfies
\[
U'_R \lesssim 2u|f| \left( \frac{n^2}{3} + \sum_{k=1}^{n} \frac{4n^2(k + 2)}{\pi^2 k^2} \right) \sim \frac{8}{\pi^2} u|f| n^2 \log n,
\]
where $u$ is the unit-roundoff, and with the assumption $nu < 1/2$.

**Proof.** If $\zeta = x_0 = 1$, the formula for $w_{k,1}$ (2.2) with $\zeta$ shifted to 0 becomes
\[
w_{k,1} S_{n-1} \left( \{0, 1 - x_1, \ldots, 1 - x_n\} - \{1 - x_k\} \right).
\]
If $k = 0$, we have
\[
w_{k,1} = \epsilon_1^0,
\]
and if $k > 1$, we have
\[
w_{k,1} = \frac{W_0}{W_1} \frac{1}{1 - z_k} \epsilon_{m-1}^{\ell,k}.
\]
To complete the proof, we may obtain asymptotics for $w_{k,1}$ as in the previous proof and use $\gamma_{k+1} \leq 2(k + 1)u$ which holds under the assumption $nu < 1/2$. 

8
Comparison of Theorems 7 and 8 gives an indication of the advantage obtained by computing the weights $w_{k,m}$ accurately followed by careful summation. Since $\gamma_n \approx nu$, the bound for the first derivative in Theorem 7 increases at the rate $n^3$. In Theorem 8, the rate is $n^2 \log n$. Thus an $n$ is replaced by $\log n$. This is very similar to the advantage obtained using compensated summation and other methods of precise summation [5]. The comparison also shows that a sound method for calculating the weights $w_{k,m}$ introduces only a modest amount of error.

3 Discretization error

In this section, we will give a discussion of the discretization error. We show that the discretization error goes up a factor of $n^2$ with every additional derivative just like the rounding error. This implies that the value of $n$ where the total error transitions from mostly due to discretization to mostly due to rounding is independent of the order of the derivative. This implication is illustrated computationally.

The Lagrange interpolant may be augmented with the remainder term as follows [2, 3]:

$$f(x) = \sum_{k=0}^{n} f(x_k) \ell_k(x) + f[x_0, x_1, \ldots, x_n, x] \prod_{k=0}^{n} (x - x_j).$$

Here the $f[]$ notation is for divided differences. The finite difference approximation to $f^{(m)}(x)$ is obtained by differentiating the Lagrange interpolant $m$ times. Therefore the discretization error for the $m$-th derivative at $x = \zeta$ is equal to

$$\frac{d^m}{d\zeta^m} f[x_0, \ldots, x_n, \zeta] \prod_{j=0}^{n} (\zeta - x_j).$$

The product rule for differentiation gives

$$\sum_{j=0}^{m} \binom{m}{j} \frac{d^j}{d\zeta^j} f[z_0, \ldots, z_n, \zeta] \prod_{k=0}^{n} (\zeta - x_k). \quad (3.1)$$

Using standard properties of divided differences, and assuming $f$ to be differentiable as many times as necessary, we may write the discretization error as

$$\sum_{j=0}^{m} m! f[x_0, \ldots, x_n, \zeta^{(j+1)}] \prod_{k=0}^{n} (\zeta - x_k),$$

where $\zeta^{(j+1)}$ stands for $\zeta$ repeated $j + 1$ times in the divided difference. Here we have used an identity for differentiating a divide difference [2]. If $x_0, \ldots, x_n$ are the Chebyshev points and the divided differences are assumed to be relatively uniform throughout the domain, this expression above shows that the discretization error too is likely to be maximum at the edges. Therefore we take $\zeta = x_0 = 1$ to get

$$U_D = \sum_{j=0}^{m} m! f[x_1, \ldots, x_n] \prod_{k=0}^{n} (1 - x_k).$$
We will denote the divided difference $f_{[1^{j+2}, x_1, \ldots, x_n]}$ by $D_{j+2}$. The expression for the discretization error becomes

$$U_D = \sum_{j=0}^{m} m! D_{j+2} S_{n+1-m+j}(1-x_1, \ldots, 1-x_n).$$  \(3.2\)

As $n$ increases, the asymptotics of $U_D$ are given by

$$U_D \sim 4 \left( \frac{D_2}{n^{2n}} \right) n^2 + \frac{8nD_3}{2^n} + \frac{24nD_4}{2^n} \sim \frac{16}{105} \left( \frac{D_2}{n^{2n}} \right) n^8 + \frac{16n^5D_3}{5.2^n} + \frac{32n^3D_4}{2^n} + \frac{96nD_5}{2^n}$$  \(3.3\)

for orders of differentiation $m = 1, 2, 3, 4$, respectively. The symmetric function $S_{n+1-m+j}$ in (3.2) is equal to $W_0 E_0^{m-j}$. From this point the symmetric functions may be estimated as in the previous section to derive (3.3).

We do not attempt to estimate the divided differences $D_{j+2}$. However they may be estimated using contour integration as shown in [10]. If $f(x) = \sin Kx$, and $K = n\pi/\eta$ with $\eta > \pi$, implying more than $\pi$ points per wavelength, the divided difference $D_2$ decreases exponentially with $n$. On the other hand, if $K$ has a fixed value such as $K = 2\pi$ the divided difference decreases super-exponentially with $n$. For functions such as $f(x) = \sin \pi x$, the divided differences $D_1, D_2$, and so on typically vary only by constant factors and the asymptotics in (3.3) may be expected to be dominated by the $D_2$ term.

The interpolation error at $x = \zeta$ is given by

$$f_{[x_1, x_2, x_3, \ldots, x_n, \zeta]} (\zeta - x_1) \ldots (\zeta - x_n).$$

Comparison with (2.7) shows that the interpolation error is approximately

$$f_{[x_1, x_2, x_3, \ldots, x_n, \zeta]} C n^{-2n} \quad (3.4)$$

for $\zeta \in (x_1, x_0)$ and where $C$ is $O(1)$.

Comparison of Theorem 7 and (3.3) suggests that the transition from discretization error to rounding error should be relatively independent of the order of the derivative. Every time the order goes up by 1, both estimates increases by a factor $n^2$, ignoring constants. Therefore the transition should be at about the same value of $n$ independently of the order of differentiation. This phenomenon is illustrated in Figure 3.1.

### 4 Choice of the mapping parameter

The choice of the parameter $\alpha$ is taken to be given by

$$\left( \frac{1 - \sqrt{1 - \alpha^2}}{\alpha} \right)^n = n^\beta u \quad (4.1)$$
Figure 3.1: Plots of error vs \( n \). The plots on the left are for \( f(x) = \sin 2\pi x \). The plots on the right are for \( f(x) = \sin Kx \) with \( K = n\pi/4 \) implying 4 points per wavelength. The solid line is the actual error. The dashed line is the discretization error computed using (3.1), with divided differences computed in extended precision. The dotted lines are the asymptotic rounding error bounds of Theorem 7. The dotted lines with circles replace \( \gamma_{6n+4-m} \) by \( nu \) and the dotted lines with squares replace that quantity by \( u \).
with $\beta = 0$ \cite{4,7} and with $u$ being the unit roundoff. We will attempt to justify this choice for all orders of derivative $m$.

Given a function $f(x)$, such as $f(x) = \sin Kx$, the mapped function is $F(\xi) = f(g(\xi))$ where $g(\cdot)$ is the mapping \cite{1.1}. We will first argue for \cite{4.1} as a balance between the discretization error and the rounding error in interpolation. The analysis of rounding error that arises in spectral differentiation is precise. The indeterminacy in the rounding error is limited to a factor of $n$, as may be seen from Figure 3.1. However, the discretization errors cannot be estimated as precisely because the divided differences that arise in (3.3) and (3.4) are not known within factors of $n$.

If $f(x) = 1$ then $F(\xi) = g(\xi) = \arcsin \alpha \xi / \arcsin \alpha$. The Chebyshev series of $F(\xi)$ may be computed from the Laurent series of $F((z + 1/z)/2)$ centered at $z = 0$ (if $z = e^{i\theta}$ then $\xi = \cos \theta$, and $(z^n + 1/z^n)/2 = \cos n\theta$ is the Chebyshev polynomial $T_n(\xi)$. If $F(\xi) = g(\xi)$, the singularities are at

$$z = \frac{\pm 1 \pm \sqrt{1 - \alpha^2}}{\alpha}.$$ 

Therefore the coefficients of $z^{\pm n}$ in the Laurent series fall off in magnitude at the rate

$$\left( \frac{1 - \sqrt{1 - \alpha^2}}{\alpha} \right)^n$$

and so does the coefficient of $T_n(\xi)$ in the Chebyshev series of $g(\xi)$. In fact, one can be more precise. Because the singularities of $g(\xi)$ at $\xi = \pm 1/\alpha$ are of the type $(\xi \mp 1/\alpha)^{1/2}$, the coefficients will fall off at the rate

$$n^{-3/2} \left( \frac{1 - \sqrt{1 - \alpha^2}}{\alpha} \right)^n. \quad (4.2)$$

This may be taken as an estimate of the discretization error in $g(\xi)$.

When $f(x) = \sin Kx$, the estimate \cite{4.2} for interpolation error will still hold but with additional modulation factors of the type $n^\beta$ with $\beta > 0$. These modulating factors are not precisely known but they certainly exist. For example, if $K = n\pi/4$, implying 4 points per wavelength, the number of terms in the expansion of $\sin K g(\xi)$ (in powers of $g(\xi)$) before the exponentially decay of coefficients kicks in, is greater than $O(n)$.

As far as the rounding error in interpolation is concerned, this quantity is bounded by $C n \log nu$, with $C$ being a small constant \cite{6}. Thus balancing of discretization and rounding errors leads to \cite{4.2} but with an indeterminacy in the exact value of $\beta$ and in constants. The appropriate balance, ignoring constants, is given by \cite{4.1}. Empirically, $\beta = 0$ is found to be a good choice although other $\beta$ such as $\beta = -1.5$ seem to do just as well. See Figure 4.1.

As long as constants are ignored, \cite{4.2} remains the right equation for balancing errors for the first derivative as well. The derivative $F'(\xi)$ is approximated by spectral differencing at the Chebyshev points. The discretization error as well as the interpolation error at the edge $\xi = 1$ go up by a factor of $n^2$ from Theorem \cite{7, 3.3}, and (3.4). The errors are pulled back into the $x$-domain through the same $g^{-1}$ transformation, and the balancing equation remains the same.

For higher derivatives $F^{(m)}(\xi)$ the balancing equation again remains the same, ignoring constants. With every increase in $m$ by 1, the discretization and rounding errors both go up
Figure 4.1: Graphs of error vs $n$ for $\sin 2\pi x$ and $\sin n\pi x/4$. The mapping parameter $\alpha$ is determined using (4.1). The errors are for the 2nd derivative.

Figure 4.2: Graphs of error vs $n$ for $\sin 2\pi x$ and $\sin n\pi x/4$. The mapping parameter $\alpha$ is determined using (4.1). The errors are for the 3rd derivative.

by a factor of $n^2$. Both errors are pulled back using the same transformation $g^{-1}$. If derivatives $f^{(m)}(x)$ are computed by successively taking the first derivative (as in [4]), rather than using a differencing scheme for the $m$-th derivative directly, the argument changes only slightly.

The discrete cosine transform is a faster method of approximating $F'(\xi)$. However, it appears to incur greater rounding error [4]. This suggests trying to balance errors in (4.1) with $\beta > 0$, as the greater error can only be due to rounding. In Figure (4.2), $\beta = 0.5$ does give smaller errors for $f(x) = \sin 2\pi x$ and the rounding errors vary more smoothly for $f(x) = \sin n\pi x/4$.

5 Acknowledgements

I am very grateful to Hans Johnston for many helpful discussions. This research was partially supported by NSF grant DMS-1115277.
References

[1] R. Baltensperger and M.R. Trummer. Spectral differencing with a twist. *SIAM J. Sci. Comp.*, 24:1465–1487, 2003.

[2] S.D. Conte and C. de Boor. *Elementary Numerical Analysis*. McGraw-Hill, 1980.

[3] Philip J. Davis. *Interpolation and Approximation*. Dover, 1975.

[4] W. S. Don and A. Solomonoff. Accuracy enhancement for higher derivatives using Chebyshev collocation and a mapping technique. *SIAM J. Sci. Comput.*, 18:1040–1055, 1997.

[5] N.J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, 2nd edition, 2002.

[6] N.J. Higham. The numerical stability of barycentric Lagrange interpolation. *IMA Journal of Numerical Analysis*, 24:547–556, 2004.

[7] D. Kosloff and H. Tal-Ezer. A modified Chebyshev pseudospectral method with an $O(N^{-1})$ time step restriction. *Journal of Computational Physics*, 104:457–469, 1993.

[8] R. Navarrete and D. Viswanath. Accuracy and stability of inversion of power series. *IMA Journal of Numerical Analysis*, 2015. in press.

[9] B. Sadiq and D. Viswanath. Finite difference weights, spectral differentiation, and superconvergence. *Mathematics of Computation*, 83:2403–2427, 2014.

[10] J.A.C. Weideman and L.N. Trefethen. The eigenvalues of second-order spectral differentiation matrices. *SIAM J. Numer. Anal.*, 25:1279–1298, 1988.