In this paper, we show that the monomial basis is generally as good as a well-conditioned polynomial basis for interpolation, provided that the condition number of the Vandermonde matrix is smaller than the reciprocal of machine epsilon. We also show that the monomial basis is more advantageous than other polynomial bases in a number of applications.

Keywords: polynomial interpolation; monomials; Vandermonde matrix; backward error analysis

Is polynomial interpolation in the monomial basis unstable?

Zewen Shen[†⋄⋆] and Kirill Serkh[‡⋄]
University of Toronto NA Technical Report
v2, February 27, 2023

[⋄] This author's work was supported in part by the NSERC Discovery Grants RGPIN-2020-06022 and DGECR-2020-00356.

[†] Dept. of Computer Science, University of Toronto, Toronto, ON M5S 2E4
[‡] Dept. of Math. and Computer Science, University of Toronto, Toronto, ON M5S 2E4

[⋆] Corresponding author

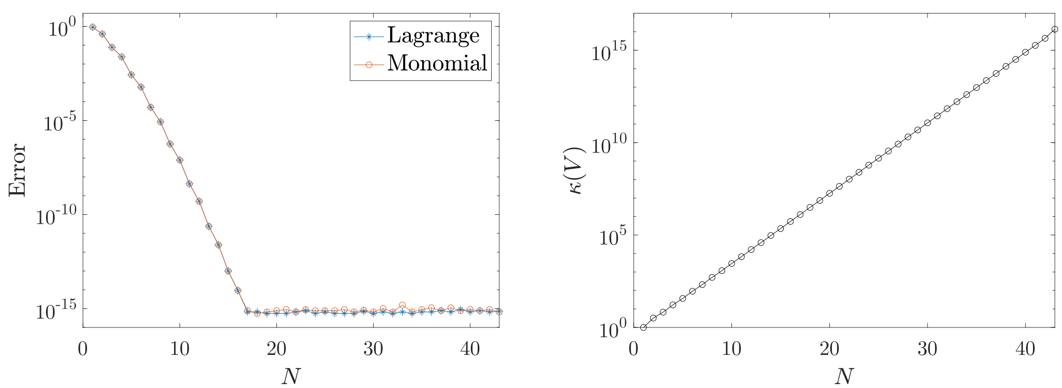
Contents

1 Introduction

Function approximation has been a central topic in numerical analysis since its inception. One of the most effective methods for approximating a function $F : [-1, 1] \to \mathbb{R}$ is the use of an interpolating polynomial P_N of degree N which satisfies $P_N(x_j) = F(x_j)$ for a set of $(N + 1)$ collocation points $\{x_j\}_{j=0,1,...,N}$. In practice, the collocation points are typically chosen to be the Chebyshev points, and the resulting interpolating polynomial, known as the Chebyshev interpolant, is a nearly optimal approximation to F in the space of polynomials of degree at most N [32]. A common basis for representing the interpolating polynomial P_N is the Lagrange polynomial basis, and the evaluation of P_N in this basis can be done stably using the Barycentric interpolation formula [8, 22]. Some other commonly used bases are Newton polynomials, Chebyshev polynomials, and Legendre polynomials. Alternatively, the monomial basis can be used to represent P_N, such that $P_N(x) = \sum_{k=0}^{N} a_k x^k$ for some coefficients $\{a_k\}_{k=0,1,...,N}$. The computation of the monomial coefficient vector $a := (a_0, a_1, \ldots, a_N)^T \in \mathbb{R}^{N+1}$ of the interpolating polynomial P_N requires the solution to a linear system $Va = f$, where

$$V := \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^N \\ 1 & x_1 & x_1^2 & \cdots & x_1^N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^N \end{pmatrix} \in \mathbb{R}^{(N+1)\times(N+1)} \tag{1}$$

is a Vandermonde matrix, and $f := \big(F(x_0), F(x_1), \ldots, F(x_N)\big)^T \in \mathbb{R}^{N+1}$ is a vector of the function values of F at the $(N + 1)$ collocation points on the interval $[-1, 1]$. It is well-known that, when all of the collocation points are real, the condition number of a Vandermonde matrix grows at least exponentially as N increases [7]. It follows that the numerical solution to this linear system is highly inaccurate when N is not small, and, as a result, this algorithm for constructing P_N is often considered to be unstable. But, is this really the case? Let $\{x_j\}_{j=0,1,\dots,N}$ be the set of $(N+1)$ Chebyshev points on the interval $[-1,1]$, and consider the case where $F(x)=\cos(2x+1)$. We solve the resulting Vandermonde system using LU factorization with partial pivoting. In Figure 1a, we present a comparison between the approximation error of the computed monomial expansion (labeled as "Monomial") and the approximation error of the Chebyshev interpolant evaluated using the Barycentric interpolation formula (labeled as "Lagrange"). One can observe that the computed monomial expansion is, surprisingly, as accurate as the Chebyshev interpolant evaluated using the Barycentric interpolation formula (which is accurate up to machine precision), despite the huge condition number of the Vandermonde matrix reported in Figure 1b.

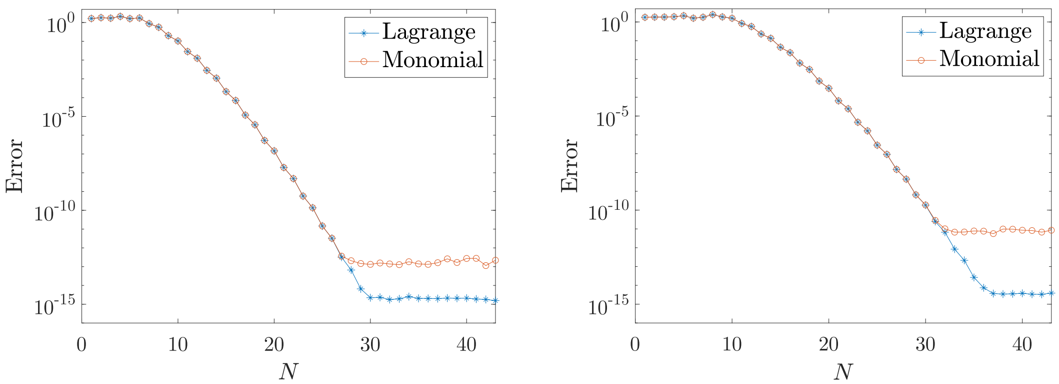


(a) L^∞ error over $[-1,1]$

(b) Condition number

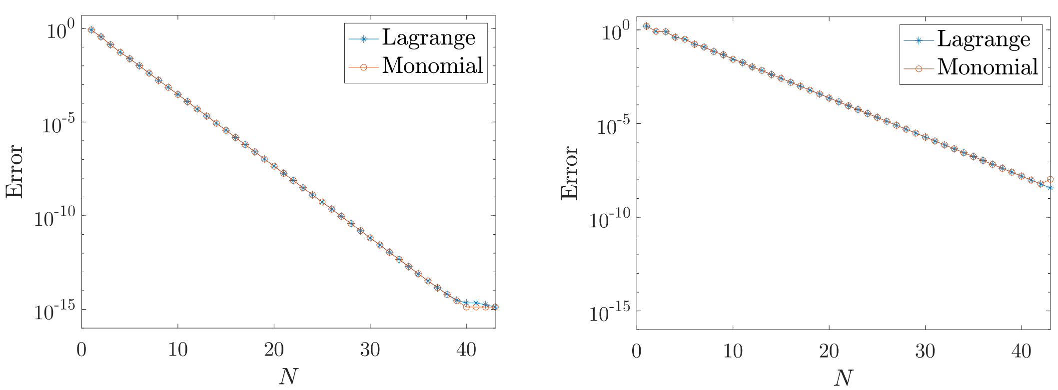
Figure 1: **Polynomial interpolation of** $\cos(2x+1)$ **in the monomial basis**. The x-axis label N denotes the order of approximation. The y-axis label "Error" denotes the L^∞ approximation error over $[-1,1]$, which is estimated by comparing the approximated function values at 10000 equidistant points over $[-1,1]$ with the true function values.

What happens when the function F becomes more complicated? In Figure 2, we compare the accuracy of the two approximations when $F(x)=\cos(8x+1)$ and when $F(x)=\cos(12x+1)$. Initially, the computed monomial expansion is as accurate as the Chebyshev interpolant evaluated using the Barycentric interpolation formula. However, the convergence of polynomial interpolation in the monomial basis stagnates after reaching a certain error threshold. Furthermore, it appears that, the more complicated a function is, the larger that error threshold becomes. But what does it mean for a function to be complicated in this context? Consider the case when the function requires an even higher-order Chebyshev interpolant in order to be approximated to machine precision. In Figure 3, we compare the accuracy of the two approximations when $F(x)=\frac{1}{x-\sqrt{2}}$ and when $F(x)=\frac{1}{x-0.5i}$. These two functions each have a singularity in a neighborhood of the interval $[-1,1]$, and Chebyshev interpolants of degree ≥ 40 are required to approximate them to machine precision. Yet, no stagnation of convergence is observed. Based on all of the previous examples, we conclude that polynomial interpolation in the monomial basis is not as unstable as it appears, and has some subtleties lurking around the corner that are worth further investigation.



(a) $\cos(8x+1)$ (b) $\cos(12x+1)$

Figure 2: **Polynomial interpolation of more complicated functions in the monomial basis**.



(a) $\frac{1}{x-\sqrt{2}}$ (b) $\frac{1}{x-0.5i}$

Figure 3: **Polynomial interpolation of functions with a singularity near the interval** $[-1,1]$ **in the monomial basis**.

These seemingly mysterious experiments can be explained partially from the point of view of backward error analysis. Indeed, the forward error $\|a-\widehat{a}\|_2$ of the numerical solution $\widehat{a}$ to the Vandermonde system $Va=f$ can be huge, but it is the backward error, i.e., $\|V\widehat{a}-f\|_2$, that matters for the accuracy of the approximation. This is because a small backward error implies that the difference between the computed monomial expansion, which we denote by $\widehat{P}_N$, and the exact interpolating polynomial, P_N, is a polynomial that approximately vanishes at all of the collocation points. When the Lebesgue constant associated with the collocation points is small (which is the case for the Chebyshev points), the polynomial $P_N-\widehat{P}_N$ is bounded uniformly by the backward error times a small constant. As a result, we bound the monomial approximation error

$\|F - \widehat{P}_N\|_{L^\infty([-1,1])}$ by the following inequality:

$$\|F - \widehat{P}_N\|_{L^\infty([-1,1])} \leq \|F - P_N\|_{L^\infty([-1,1])} + \|P_N - \widehat{P}_N\|_{L^\infty([-1,1])}. \tag{2}$$

We refer to the first and the second terms on the right-hand side of (2) as the polynomial interpolation error and the backward error, respectively. When the backward error is smaller than the polynomial interpolation error, the monomial approximation error is dominated by the polynomial interpolation error, and the use of a monomial basis does not incur any additional loss of accuracy. Once the polynomial interpolation error becomes smaller than the backward error, the convergence of the approximation stagnates. For example, in Figure 3a, we verify numerically that the backward error is around the size of machine epsilon for all $N \leq 43$, so stagnation is not observed, and polynomial interpolation in the monomial basis is as accurate as polynomial interpolation in the Lagrange basis, evaluated by the Barycentric interpolation formula. On the other hand, in Figure 2a, the backward error is around the size of 10^{-13} for $N \geq 20$, which leads to stagnation once the polynomial interpolation error is less than 10^{-13}.

The explanation above brings up a new question: when will the backward error be small? When a backward stable linear system solver (e.g., LU factorization with partial pivoting) is used to solve the Vandermonde system $Va = f$, it is guaranteed that the numerical solution $\widehat{a}$ is the exact solution to the linear system

$$(V + \delta V)\widehat{a} = f, \tag{3}$$

for a matrix $\delta V \in \mathbb{R}^{(N+1)\times(N+1)}$ that satisfies $\|\delta V\|_2 \leq u \cdot \gamma$, where u denotes machine epsilon and $\gamma = \mathcal{O}(\|V\|_2)$. It follows that the backward error, $\|V\widehat{a} - f\|_2$, of the numerical solution is bounded by $u \cdot \gamma \|\widehat{a}\|_2$. We note that γ is typically small, so the backward error is essentially determined by the norm of the computed monomial coefficient vector. In fact, so long as $\kappa(V) \lesssim \frac{1}{u}$, one can show that the norm of the monomial coefficient vector computed by a backward stable solver is around the same size as the norm of the exact monomial coefficient vector of the interpolating polynomial. Therefore, in this case, the monomial approximation error can be quantified a priori using information about the interpolating polynomial, which implies that a theory of polynomial interpolation in the monomial basis can be developed.

The rest of the paper is organized as follows. In Section 2, we analyze polynomial interpolation in the monomial basis over a smooth simple arc in the complex plane, with the interval as a special case, along with a number of numerical experiments. Our analysis provides insight into the suitability and limitations of the monomial basis as a viable alternative to well-conditioned polynomial bases for interpolation. In Section 3, we present applications where the use of a monomial basis for interpolation offers a substantial advantage over other bases. In Section 4, we review related work, and discuss the generalization of our theory to higher dimensions.

2 Polynomial interpolation in the monomial basis

Let $\Gamma \subset \mathbb{C}$ be a smooth simple arc, and let $F : \Gamma \to \mathbb{C}$ be an arbitrary function. The Nth degree interpolating polynomial, denoted by P_N, of the function F for a given set of $(N+1)$ distinct collocation points $Z := \{z_j\}_{j=0,1,\ldots,N} \subset \Gamma$ can be expressed as $P_N(z) = \sum_{k=0}^{N} a_k z^k$, where the monomial coefficient vector $(a_0, a_1, \ldots, a_N)^T$ is the solution to the Vandermonde system

$$\begin{pmatrix} 1 & z_0 & z_0^2 & \cdots & z_0^N \\ 1 & z_1 & z_1^2 & \cdots & z_1^N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & z_N & z_N^2 & \cdots & z_N^N \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} F(z_0) \\ F(z_1) \\ \vdots \\ F(z_N) \end{pmatrix}. \tag{4}$$

For ease of notation, we denote the Vandermonde matrix by $V^{(N)}$, the monomial coefficient vector by $a^{(N)}$, and the corresponding right-hand side vector by $f^{(N)}$.

In order to study the size of the residual of the numerical solution to the Vandermonde system, we require the following lemma, which provides a bound for the 2-norm of the solution to a perturbed linear system.

Lemma 2.1. *Let N be a positive integer. Suppose that $A \in \mathbb{C}^{N\times N}$ is invertible, $b \in \mathbb{C}^N$, and that $x \in \mathbb{C}^N$ satisfies $Ax = b$. Suppose further that $\widehat{x} \in \mathbb{C}^N$ satisfies $(A + \delta A)\widehat{x} = b$ for some $\delta A \in \mathbb{C}^{N\times N}$. If there exists an $\alpha > 1$ such that*

$$\|A^{-1}\|_2 \le \frac{1}{\alpha \cdot \|\delta A\|_2}, \tag{5}$$

then the matrix $A + \delta A$ is invertible, and $\widehat{x}$ satisfies

$$\frac{\alpha}{\alpha+1}\|x\|_2 \le \|\widehat{x}\|_2 \le \frac{\alpha}{\alpha-1}\|x\|_2. \tag{6}$$

Proof. By multiplying both sides of $(A + \delta A)\widehat{x} = b$ by A^{-1}, we have that

$$(I + A^{-1}\delta A)\widehat{x} = x, \tag{7}$$

where I denotes the identity matrix. By (5), the term $A^{-1}\delta A$ satisfies

$$\|A^{-1}\delta A\|_2 \le \|A^{-1}\|_2\|\delta A\|_2 \le \frac{1}{\alpha} < 1. \tag{8}$$

Thus, it follows that the matrix $A + \delta A$ is invertible, and $\|\widehat{x}\|_2$ satisfies

$$\|\widehat{x}\|_2 \le \|(I + A^{-1}\delta A)^{-1}\|_2\|x\|_2 \le \frac{1}{1 - \|A^{-1}\delta A\|_2}\|x\|_2 \le \frac{\alpha}{\alpha-1}\|x\|_2. \tag{9}$$

In addition, by (8), $\|x\|_2$ satisfies

$$\|x\|_2 \le \|I + A^{-1}\delta A\|_2\|\widehat{x}\|_2 \le \left(1 + \frac{1}{\alpha}\right)\|\widehat{x}\|_2. \tag{10}$$

The proof is complete by combining (9) and (10). ■

The following theorem provides an a priori upper bound for the monomial approximation error.

Theorem 2.2. *Let $\Gamma \subset \mathbb{C}$ be a smooth simple arc, and let $F : \Gamma \to \mathbb{C}$ be an arbitrary function. Suppose that P_N is the Nth degree interpolating polynomial of F for a given set of $(N+1)$ distinct collocation points $Z := \{z_j\}_{j=0,1,\dots,N} \subset \Gamma$. Clearly, the monomial coefficient vector $a^{(N)}$ of the polynomial P_N is the solution to the Vandermonde system $V^{(N)}a^{(N)} = f^{(N)}$, where $V^{(N)}$ and $f^{(N)}$ have been previously defined in (4). Suppose further that there exists some constant $\gamma_N \geq 0$ such that the computed monomial coefficient vector $\widehat{a}^{(N)} = (\widehat{a}_0, \widehat{a}_1, \dots, \widehat{a}_N)^T$ satisfies*

$$\big(V^{(N)} + \delta V^{(N)}\big)\widehat{a}^{(N)} = f^{(N)}, \tag{11}$$

for some $\delta V^{(N)} \in \mathbb{C}^{(N+1)\times(N+1)}$ with

$$\|\delta V^{(N)}\|_2 \leq u \cdot \gamma_N, \tag{12}$$

where u denotes machine epsilon. Let $\widehat{P}_N(z) := \sum_{k=0}^{N} \widehat{a}_k z^k$ be the computed monomial expansion. If the 2-norm of $(V^{(N)})^{-1}$ satisfies

$$\|(V^{(N)})^{-1}\|_2 \leq \frac{1}{2u \cdot \gamma_N}, \tag{13}$$

then the 2-norm of the numerical solution $\widehat{a}^{(N)}$ is bounded by

$$\frac{2}{3}\|a^{(N)}\|_2 \leq \|\widehat{a}^{(N)}\|_2 \leq 2\|a^{(N)}\|_2, \tag{14}$$

and the monomial approximation error can be quantified a priori by

$$\|F - \widehat{P}_N\|_{L^\infty(\Gamma)} \leq \|F - P_N\|_{L^\infty(\Gamma)} + 2u \cdot \gamma_N \Lambda_N \|a^{(N)}\|_2, \tag{15}$$

where Λ_N denotes the Lebesgue constant for Z.

Proof. By combining (12) and (13), we have that

$$\|\delta V^{(N)}\|_2 \leq \frac{1}{2 \cdot \|(V^{(N)})^{-1}\|_2}. \tag{16}$$

It follows from Lemma 2.1 that the 2-norm of the computed monomial coefficient vector $\widehat{a}^{(N)}$ is bounded by

$$\frac{2}{3}\|a^{(N)}\|_2 \leq \|\widehat{a}^{(N)}\|_2 \leq 2\|a^{(N)}\|_2, \tag{17}$$

Consequently, the backward error of the numerical solution to (11) satisfies

$$\|V^{(N)}\widehat{a}^{(N)} - f^{(N)}\|_2 \leq u \cdot \gamma_N \|\widehat{a}^{(N)}\|_2 \leq 2u \cdot \gamma_N \|a^{(N)}\|_2. \tag{18}$$

Furthermore, by the triangle inequality, the definition of the Lebesgue constant Λ_N, and (18), the monomial approximation error satisfies

$$\begin{aligned}
\|F - \widehat{P}_N\|_{L^\infty(\Gamma)} &\leq \|F - P_N\|_{L^\infty(\Gamma)} + \|\widehat{P}_N - P_N\|_{L^\infty(\Gamma)} \\
&\leq \|F - P_N\|_{L^\infty(\Gamma)} + \Lambda_N \|V^{(N)}\widehat{a}^{(N)} - f^{(N)}\|_2 \\
&\leq \|F - P_N\|_{L^\infty(\Gamma)} + 2u \cdot \gamma_N \Lambda_N \|a^{(N)}\|_2. \qquad (19)
\end{aligned}$$

■

When the Vandermonde system is solved by a backward stable linear system solver, the set of assumptions (11) and (12) is satisfied with constant $\gamma_N = \mathcal{O}(\|V^{(N)}\|_2)$, from which it follows that the precondition (13) becomes $\kappa(V^{(N)}) \lesssim \frac{1}{u}$. Without loss of generality, one can assume that Γ is inside the unit disk D_1 centered at the origin, such that $\|V^{(N)}\|_2$ is small. In this case, we observe that $\gamma_N \lesssim 1$ for at least $N \leq 100$ when LU factorization with partial pivoting (which is backward stable) is used to solve the Vandermonde system.

Note that the second term on the right-hand side of (15) is an upper bound of the backward error $\|P_N - \widehat{P}_N\|_{L^\infty(\Gamma)}$, i.e., the extra loss of accuracy caused by the use of a monomial basis. Additionally, the absolute condition number of the evaluation of $P_N(z)$ in the monomial basis is around $\|a^{(N)}\|_2$ when $|z| \approx 1$. Therefore, even if the monomial coefficients of P_N were known analytically, the cancellation error associated with the evaluation of P_N in the monomial basis over Γ is expected to be of comparable magnitude to this upper bound of the backward error, provided that: (1) the Lebesgue constant Λ_N is not large; (2) a backward stable linear system solver is used to solve the Vandermonde system; and (3) $\|(V^{(N)})^{-1}\|_2 \leq \frac{1}{2u\cdot\gamma_N}$. Meeting the first two conditions is straightforward, since any well-chosen set of collocation points results in a small Λ_N, and essentially any standard linear system solver is backward stable. The third condition requires further attention, and we address it in Section 2.2.

The rest of this section is structured as follows. First, we review a classical result on function approximation over a smooth simple arc $\Gamma \subset \mathbb{C}$ by polynomials. Next, we study the backward error $\|P_N - \widehat{P}_N\|_{L^\infty(\Gamma)}$ by bounding the 2-norm of the monomial coefficients of the interpolating polynomial. Finally, we study the growth of $\|(V^{(N)})^{-1}\|_2$, which determines the validity of the precondition on the a priori error estimate (15).

Below, we define a generalization of the Bernstein ellipse, to the case of a smooth simple arc in the complex plane.

Definition 2.1. Given a smooth simple arc Γ in the complex plane, we define E_ρ to be the level set $\{x + iy \in \mathbb{C} : G(x, y) = \log \rho\}$, where $G : \mathbb{R}^2 \to \mathbb{R}$ is the unique solution to the exterior Laplace equation

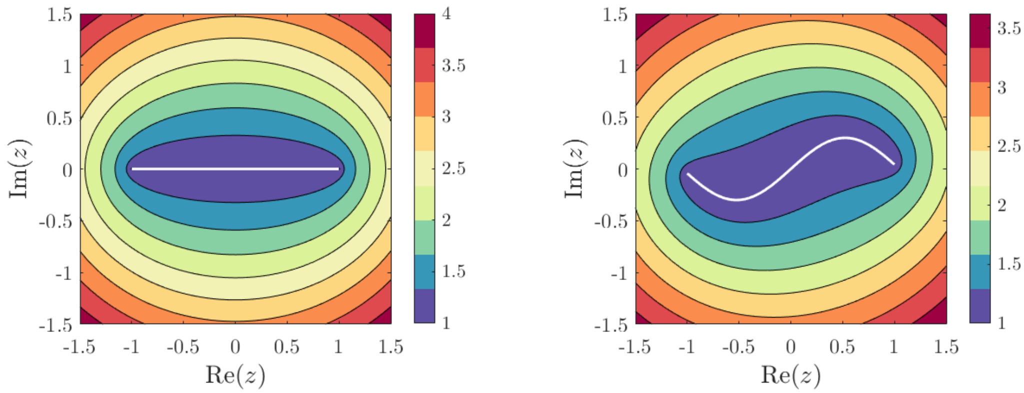
$$\begin{aligned} \nabla^2 G &= 0 \text{ in } \mathbb{R}^2 \setminus \Gamma, \\ G &= 0 \text{ on } \partial\Gamma, \\ G(x) &\sim \log|x| \text{ as } |x| \to \infty. \end{aligned} \tag{20}$$

Furthermore, we let E_ρ^o denote the open region bounded by E_ρ.

We note that, when $\Gamma = [a, b] \subset \mathbb{R}$, the level set E_ρ is a Bernstein ellipse with parameter ρ, with foci at a and b. In Figure 4, we plot examples of level sets E_ρ for an interval and for a sine curve, for various values of ρ.

The following lemma demonstrates the feasibility of function approximation by polynomials over a smooth simple arc Γ in the complex plane. We refer the readers to Section 4.5 in [33] for the proof.

Lemma 2.3. *Let Γ be a smooth simple arc in the complex plane. Suppose that the function $F : \Gamma \to \mathbb{C}$ is analytically continuable to the closure of the region E_ρ^o corresponding to Γ,*



(a) $\Gamma = [-1, 1]$

(b) $\Gamma = \{t + 0.3i\sin(3t) : t \in [-1, 1]\}$

Figure 4: **The level set E_ρ corresponding to Γ, for various values of ρ.** The colorbar indicates the value of ρ. The smooth simple arc Γ is the white curve in the figure. The plots were made using the source code provided in [2].

for some $\rho > 1$. Then, there exists a sequence of polynomials $\{Q_n\}$ satisfying

$$\|F - Q_n\|_{L^\infty(\Gamma)} \le C\rho^{-n}, \tag{21}$$

for all $n \ge 0$, where $C \ge 0$ is a constant that is independent of N.

Remark 2.1. When Γ is a line segment, the magnitude of the constant C in (21) is proportional to $\|F\|_{L^\infty(E_\rho^o)}$ (see Lemma 2.8 in Section 2.3). We conjecture that the same holds in the general case.

The parameter ρ_* defined below appears in our bounds for both the 2-norm of the monomial coefficient vector of the interpolating polynomial, and the growth rate of the 2-norm of the inverse of a Vandermonde matrix. It denotes the parameter of the smallest region E_ρ^o that contains the open unit disk centered at the origin.

Definition 2.2. Given a smooth simple arc $\Gamma \subset \mathbb{C}$, define $\rho_* := \inf\{\rho > 1 : D_1 \subset E_\rho^o\}$, where D_1 is the open unit disk centered at the origin, and E_ρ^o is the region corresponding to Γ (see Definition 2.1).

The following lemma provides upper bounds for the 2-norm of the monomial coefficient vector of an arbitrary polynomial.

Lemma 2.4. *Let $P_N : \mathbb{C} \to \mathbb{C}$ be a polynomial of degree N, where $P_N(z) = \sum_{k=0}^{N} a_k z^k$ for some $a_0, a_1, \dots, a_N \in \mathbb{C}$. The 2-norm of the coefficient vector $a^{(N)} := (a_0, a_1, \dots, a_N)^T$ satisfies*

$$\|a^{(N)}\|_2 \le \|P_N\|_{L^\infty(\partial D_1)} \le \rho_*^N \|P_N\|_{L^\infty(\Gamma)}, \tag{22}$$

where D_1 denotes the open unit disk centered at the origin, and ρ_* is given in Definition 2.2.

Proof. Observe that

$$P_N(e^{i\theta}) = \sum_{k=0}^{N} a_k e^{ik\theta}. \tag{23}$$

By Parseval's identity, we have that

$$\|a^{(N)}\|_2 = \Big(\frac{1}{2\pi}\int_0^{2\pi} |P_N(e^{i\theta})|^2 \, \mathrm{d}\theta\Big)^{1/2} \le \|P_N\|_{L^\infty(\partial D_1)} \le \|P_N\|_{L^\infty(E^o_{\rho_*})}, \quad (24)$$

where the last inequality comes from the fact that $D_1 \subset E^o_{\rho_*}$ (see Definition 2.1). Finally, based on one of Bernstein's inequalities (see Section 4.6 in [33]), we have that

$$\|P_N\|_{L^\infty(E^o_{\rho_*})} \le \rho_*^N \|P_N\|_{L^\infty(\Gamma)}. \quad (25)$$

■

The following theorem provides an upper bound for the 2-norm of the monomial coefficients of an arbitrary interpolating polynomial.

Theorem 2.5. *Let Γ be a smooth simple arc in the complex plane, and let $F : \Gamma \to \mathbb{C}$ be an arbitrary function. Suppose that there exists a finite sequence of polynomials $\{Q_n\}_{n=0,1,\dots,N}$, where Q_n has degree n, which satisfies*

$$\|F - Q_n\|_{L^\infty(\Gamma)} \le C_N \rho^{-n}, \quad 0 \le n \le N, \quad (26)$$

for some constants $\rho > 1$ and $C_N \ge 0$. Define $P_N(z) = \sum_{k=0}^N a_k z^k$ to be the Nth degree interpolating polynomial of F for a given set of distinct collocation points $Z = \{z_j\}_{j=0,1,\dots,N} \subset \Gamma$. The 2-norm of the monomial coefficient vector $a^{(N)} := (a_0, a_1, \dots, a_N)^T$ of P_N satisfies

$$\|a^{(N)}\|_2 \le \|F\|_{L^\infty(\Gamma)} + C_N\Big(\Lambda_N\Big(\frac{\rho_*}{\rho}\Big)^N + 2\rho_* \sum_{j=0}^{N-1}\Big(\frac{\rho_*}{\rho}\Big)^j\Big), \quad (27)$$

where ρ_* is given in Definition 2.2, and Λ_N denotes the Lebesgue constant for Z.

Proof. Given $n \ge 0$, let $M^{(n)} : \mathbb{R}^{n+1} \to \mathcal{P}_n$ be the bijective linear map associating each vector $(u_0, u_1, \dots, u_n)^T \in \mathbb{R}^{n+1}$ with the nth degree polynomial $\sum_{k=0}^n u_k z^k \in \mathcal{P}_n$. It follows immediately from Lemma 2.4 that, given any polynomial $P \in \mathcal{P}_n$,

$$\big\|(M^{(n)})^{-1}[P]\big\|_2 \le \rho_*^n \|P\|_{L^\infty(\Gamma)}. \quad (28)$$

Therefore, by the triangle inequality, the 2-norm of the monomial coefficient vector of the polynomial Q_N satisfies

$$\begin{aligned}
\big\|(M^{(N)})^{-1}[Q_N]\big\|_2 &\le \big\|(M^{(N)})^{-1}[Q_0]\big\|_2 + \sum_{j=0}^{N-1} \big\|(M^{(N)})^{-1}[Q_{j+1} - Q_j]\big\|_2 \\
&= \big\|(M^{(0)})^{-1}[Q_0]\big\|_2 + \sum_{j=0}^{N-1} \big\|(M^{(j+1)})^{-1}[Q_{j+1} - Q_j]\big\|_2 \\
&\le \|F\|_{L^\infty(\Gamma)} + \sum_{j=0}^{N-1} \rho_*^{j+1} \|Q_{j+1} - Q_j\|_{L^\infty(\Gamma)} \\
&\le \|F\|_{L^\infty(\Gamma)} + 2C_N\rho_* \sum_{j=0}^{N-1}\Big(\frac{\rho_*}{\rho}\Big)^j, \quad (29)
\end{aligned}$$

from which it follows that $\|a^{(N)}\|_2$ satisfies

$$\begin{aligned}\|a^{(N)}\|_2 &\le \left\|(M^{(N)})^{-1}[P_N - Q_N]\right\|_2 + \left\|(M^{(N)})^{-1}[Q_N]\right\|_2 \\ &\le \rho_*^N \|P_N - Q_N\|_{L^\infty(\Gamma)} + \left\|(M^{(N)})^{-1}[Q_N]\right\|_2 \\ &\le \rho_*^N \Lambda_N \|F - Q_N\|_{L^\infty(\Gamma)} + \left\|(M^{(N)})^{-1}[Q_N]\right\|_2 \\ &\le \|F\|_{L^\infty(\Gamma)} + C_N\left(\Lambda_N\left(\frac{\rho_*}{\rho}\right)^N + 2\rho_*\sum_{j=0}^{N-1}\left(\frac{\rho_*}{\rho}\right)^j\right),\end{aligned} \tag{30}$$

where the third inequality comes from the observation that $P_N - Q_N$ is the interpolating polynomial of $F - Q_N$ for the set of collocation points Z. ■

Remark 2.2. The assumption (26) made in the theorem above can be satisfied for any function F by choosing C_N to be sufficiently large.

The following theorem bounds the growth of the 2-norm of the inverse of a Vandermonde matrix.

Theorem 2.6. *Suppose that $V^{(N)} \in \mathbb{C}^{(N+1)\times(N+1)}$ is a Vandermonde matrix with $(N+1)$ distinct collocation points $Z = \{z_j\}_{j=0,1,\ldots,N} \subset \mathbb{C}$. Suppose further that $\Gamma \subset \mathbb{C}$ is a smooth simple arc such that $Z \subset \Gamma$. The 2-norm of $(V^{(N)})^{-1}$ is bounded by*

$$\|(V^{(N)})^{-1}\|_2 \le \rho_*^N \Lambda_N, \tag{31}$$

where ρ_* is given in Definition 2.2, and Λ_N denotes the Lebesgue constant for the set of collocation points Z over Γ.

Proof. Let $f^{(N)} = (f_0, f_1, \ldots, f_N)^T \in \mathbb{C}^{N+1}$ be an arbitrary vector. Suppose that P_N is an interpolating polynomial of degree N for the set $\{(z_j, f_j)\}_{j=0,1,\ldots,N}$. By Lemma 2.4, the 2-norm of the monomial coefficient vector $a^{(N)}$ of P_N satisfies

$$\|a^{(N)}\|_2 \le \rho_*^N \|P_N\|_{L^\infty(\Gamma)} \le \rho_*^N \Lambda_N \|f^{(N)}\|_\infty \le \rho_*^N \Lambda_N \|f^{(N)}\|_2, \tag{32}$$

where the second inequality follows from the definition of the Lebesgue constant. Therefore, the 2-norm of $(V^{(N)})^{-1}$ is bounded by

$$\|(V^{(N)})^{-1}\|_2 = \sup_{f^{(N)}\neq 0}\left\{\frac{\|(V^{(N)})^{-1} f^{(N)}\|_2}{\|f^{(N)}\|_2}\right\} = \sup_{f^{(N)}\neq 0}\left\{\frac{\|a^{(N)}\|_2}{\|f^{(N)}\|_2}\right\} \le \rho_*^N \Lambda_N. \tag{33}$$

■

Note that the bound above applies to any smooth simple arc $\Gamma \subset \mathbb{C}$ that contains the set of collocation points Z.

Observation 2.3. In the case where the set of $(N+1)$ collocation points $Z \subset \Gamma$ are chosen such that the associated Lebesgue constant Λ_N is small, we observe in practice that the upper bound $\rho_*^N \Lambda_N$ is reasonably close to the value of $\|(V^{(N)})^{-1}\|_2$ (see Figures 5 and 10b for numerical evidence).

2.1 When is the monomial basis as good as a well-conditioned polynomial basis?

Without loss of generality, we assume that the smooth simple arc Γ is inside the unit disk centered at the origin (such that $\|V^{(N)}\|_2$ is small and $\gamma_N \lesssim 1$), and that $F : \Gamma \to \mathbb{C}$ satisfies $\|F\|_{L^\infty(\Gamma)} \leq 1$. Furthermore, we choose a set of $(N+1)$ collocation points $Z \subset \Gamma$ with a small Lebesgue constant Λ_N, and let $V^{(N)}$ denote the corresponding Vandermonde matrix. Recall from Theorem 2.2 that, if

$$\|(V^{(N)})^{-1}\|_2 \leq \frac{1}{2u \cdot \gamma_N}, \tag{34}$$

then the monomial approximation error $\|F - \widehat{P}_N\|_{L^\infty(\Gamma)}$ is bounded a priori by

$$\|F - \widehat{P}_N\|_{L^\infty(\Gamma)} \lesssim \|F - P_N\|_{L^\infty(\Gamma)} + u \cdot \|a^{(N)}\|_2, \tag{35}$$

where u denotes machine epsilon, $\widehat{P}_N$ is the computed monomial expansion, P_N is the exact Nth degree interpolating polynomial of F for the set of collocation points Z, and $a^{(N)}$ is the monomial coefficient vector of P_N.

By Theorem 2.5, if there exists a constant $C_N \geq 0$ and a finite sequence of polynomials $\{Q_n\}_{n=0,1,\dots,N}$ such that $\|F - Q_n\|_{L^\infty(\Gamma)} \leq C_N \rho_*^{-n}$ for $0 \leq n \leq N$, where Q_n has degree n and ρ_* is given in Definition 2.2, then the monomial coefficient vector $a^{(N)}$ of P_N satisfies

$$\|a^{(N)}\|_2 \lesssim C_N \Lambda_N N \approx C_N N, \tag{36}$$

and inequality (35) becomes

$$\|F - \widehat{P}_N\|_{L^\infty(\Gamma)} \lesssim \|F - P_N\|_{L^\infty(\Gamma)} + u \cdot C_N N. \tag{37}$$

In practice, one can take $\{Q_n\}_{n=0,1,\dots,N}$ to be a finite sequence of interpolating polynomials $\{P_n\}_{n=0,1,\dots,N}$ of F for sets of collocation points with small Lebesgue constants, and define C_N by the formula

$$C_N = \min\{C \geq 0 : \|F - P_n\|_{L^\infty(\Gamma)} \leq C\rho_*^{-n} \text{ for } 0 \leq n \leq N\}. \tag{38}$$

In the case where $C_N \lesssim 1$, the inequality above shows that the extra error caused by the use of a monomial basis is around machine epsilon in size. Furthermore, by Theorem 2.6, we have that in this case $\|(V^{(N)})^{-1}\|_2 \lesssim \rho_*^N$, from which it follows that the interpolation error $\|F - P_N\|_{L^\infty(\Gamma)}$ attains a size of approximately machine epsilon before $\|(V^{(N)})^{-1}\|_2$ reaches $\frac{1}{2u\cdot\gamma_N}$. Thus, the precondition $\|(V^{(N)})^{-1}\|_2 \leq \frac{1}{2u\cdot\gamma_N}$ on the a priori error bound does not weaken the aforementioned result.

Remark 2.4. By Lemma 2.3 , if the function $F : \Gamma \to \mathbb{C}$ is analytically continuable to the closure of the region $E^o_{\rho_*}$ corresponding to Γ, then there exists a finite sequence of polynomials $\{Q_n\}_{n=0,1,\dots,N}$, where Q_n has degree n, such that $\|F - Q_n\|_{L^\infty(\Gamma)} \leq C\rho_*^{-n}$ for some $C \geq 0$. We conjecture that $C \lesssim 1$ when $\|F\|_{L^\infty(E^o_{\rho_*})} \lesssim 1$ (see also Remark 2.1).

What happens if the polynomial interpolation error decays slowly? Suppose that the polynomial interpolation error $\|F - P_n\|_{L^\infty(\Gamma)}$ decays to the value $\|F - P_N\|_{L^\infty(\Gamma)}$ at a slower rate than ρ_*^{-n}, i.e.,

$$\|F - P_n\|_{L^\infty(\Gamma)} \leq \rho_*^{N-n}\|F - P_N\|_{L^\infty(\Gamma)}, \tag{39}$$

for $0 \leq n \leq N$. In this case, $C_N = \rho_*^N\|F - P_N\|_{L^\infty(\Gamma)}$, so inequality (37) becomes

$$\|F - \widehat{P}_N\|_{L^\infty(\Gamma)} \lesssim (1 + u \cdot N\rho_*^N)\|F - P_N\|_{L^\infty(\Gamma)}. \tag{40}$$

When $N\rho_*^N \leq \frac{1}{u}$, we have that $\|F - \widehat{P}_N\|_{L^\infty(\Gamma)} \lesssim 2\|F - P_N\|_{L^\infty(\Gamma)}$. Furthermore, based on Observation 2.3, the condition $N\rho_*^N \leq \frac{1}{u}$ is generally true when $\|(V^{(N)})^{-1}\|_2 \leq \frac{1}{2u\cdot\gamma_N}$. Therefore, the extra error caused by the use of a monomial basis, despite being large in absolute terms, is less than the interpolation error $\|F - P_N\|_{L^\infty(\Gamma)}$, up until the order N reaches the threshold value corresponding to the condition $\|(V^{(N)})^{-1}\|_2 \leq \frac{1}{2u\cdot\gamma_N}$. It follows that polynomial interpolation in the monomial basis is, once again, no worse than polynomial interpolation in a well-conditioned basis in this case.

Remark 2.5. The assumption (39) holds for any sufficiently badly behaved function, e.g., functions which are not in $C^\infty(\Gamma)$, meromorphic functions with singularities close to Γ, and highly oscillatory analytic functions.

Stagnation of convergence, as shown in Figure 2, is observed when the interpolation error $\|F - P_N\|_{L^\infty(\Gamma)}$ becomes smaller than approximately $u \cdot \|a^{(N)}\|_2$. This occurs when the interpolation error $\|F - P_n\|_{L^\infty(\Gamma)}$ decays at a rate faster than ρ_*^{-n}, after previously decaying at a rate slower than ρ_*^{-n}. Suppose, as is often the case, that the function F is analytically continuable to the closure of the region E_ρ^o corresponding to Γ (see Definition 2.1), for some $\rho > \rho_*$. In this case, a generalization of Lemma 2.3 (see, for example, [10]) shows that $\|F - P_N\|_{L^\infty(E_{\rho_*}^o)} = \mathcal{O}\big((\rho_*/\rho)^N\big)$. When N is sufficiently large that $\|F - P_N\|_{L^\infty(E_{\rho_*}^o)} \leq \|F\|_{L^\infty(\partial D_1)}$, we can bound the 2-norm of the monomial coefficient vector $a^{(N)}$ of P_N by

$$\|a^{(N)}\|_2 \leq \|P_N\|_{L^\infty(\partial D_1)} \leq \|P_N - F\|_{L^\infty(\partial D_1)} + \|F\|_{L^\infty(\partial D_1)} \leq 2\|F\|_{L^\infty(\partial D_1)}, \tag{41}$$

where the first inequality comes from Lemma 2.4. Therefore, based on inequality (35), the monomial basis is at least as good as a well-conditioned polynomial basis for interpolation, until $\|F - P_N\|_{L^\infty(\Gamma)}$ decays to approximately $u \cdot \|F\|_{L^\infty(\partial D_1)}$. When there is no information available about $\|F\|_{L^\infty(\partial D_1)}$, the value of $C_N N$ can serve as a rough estimate for $\|a^{(N)}\|_2$ by inequality (36), as its calculation only involves the quantities $\{\|F - P_n\|_{L^\infty(\Gamma)}\}_{n=0,1,\dots,N}$. An even cruder bound is provided by inequality (32) in the proof of Theorem 2.6, i.e., $\|a^{(N)}\|_2 \leq \rho_*^N \Lambda_N \|f^{(N)}\|_2$, which describes the largest possible value of $\|a^{(N)}\|_2$, as the right-hand side of this inequality is independent of the smoothness of the function F. Finally, it is worth noting that, by inequality (14), $\|a^{(N)}\|_2$ can be easily estimated a posteriori using the value of $\|\widehat{a}^{(N)}\|_2$. Therefore, the extra error caused by the use of a monomial basis can always be estimated promptly during computation.

2.2 How restrictive is the monomial basis?

What are the restrictions on polynomial interpolation in the monomial basis? Firstly, extremely high-order global interpolation is impossible in the monomial basis, because the order N must satisfy $\|(V^{(N)})^{-1}\|_2 \leq \frac{1}{2u\cdot\gamma_N}$ for our estimates to hold. In fact, even if this condition were not required, there would still be no benefit to taking an order larger than this threshold in almost all situations. Suppose that the function F can only be approximated by a very high-degree interpolating polynomial to the desired accuracy. Since, in practice, $\|(V^{(N)})^{-1}\|_2 \approx \rho_*^N$ (see Observation 2.3), inequality (40) shows that the error caused by the use of the monomial basis typically dominates the approximation error whenever $\|(V^{(N)})^{-1}\|_2 > \frac{1}{2u\cdot\gamma_N}$. On the other hand, piecewise polynomial interpolation in the monomial basis over a partition of Γ can be carried out stably, provided that the maximum order of approximation over each subpanel is maintained below the threshold, and that the size of $u \cdot \|a^{(N)}\|_2 \approx u \cdot \|\widehat{a}^{(N)}\|_2$ is kept below the size of the polynomial interpolation error, where $a^{(N)}$ and $\widehat{a}^{(N)}$ denote the exact and the computed monomial coefficient vectors, respectively. As demonstrated in Section 2.1, the latter requirement is often satisfied automatically, and when it is not, adding an extra level of subdivision almost always resolves the issue. Moreover, since the convergence rate of piecewise polynomial approximation is $\mathcal{O}(h^{N+1})$, where h and N denote the maximum diameter and minimum order of approximation over all subpanels, respectively, and since the aforementioned threshold is generally not small (e.g., the threshold is approximately 43 when $\Gamma = [-1, 1]$), piecewise polynomial interpolation in the monomial basis converges rapidly so long as we set the value of N to be large enough. Therefore, there is no need to avoid the use of a monomial basis when it offers a distinct advantage over other bases.

Remark 2.6. It takes $\mathcal{O}(N^3)$ operations to solve a Vandermonde system of size $N \times N$ by a standard backward stable solver, e.g., LU factorization with partial pivoting. Since the order of approximation N is almost always not large, the solution to the Vandermonde matrix can be computed accurately, in the sense that γ_N is small, and rapidly, using highly optimized linear algebra libraries, e.g., LAPACK. There also exist specialized algorithms that solve Vandermonde systems in $\mathcal{O}(N^2)$ operations, e.g., the Björck-Pereyra algorithm [9], the Parker-Traub algorithm [16].

Observation 2.7. What happens when the order of approximation exceeds the threshold? We observe that, despite that our theory is no longer applicable, the monomial approximation error does not become much larger than the error at the threshold, when the columns of the Vandermonde matrix are ordered as in (4) and when the system is solved by MATLAB's backslash operator (which implements LU factorization with partial pivoting).

2.3 Interpolation over an interval

In this section, we consider polynomial interpolation in the monomial basis over an interval $\Gamma = [a, b] \subset \mathbb{R}$. We suggest the use of the Chebyshev points on the interval $[a, b]$ as the collocation points, because of the following two well-known lemmas related to Chebyshev approximation.

The lemma below, originally proved in [12], bounds the growth rate of the Lebesgue constant for the Chebyshev points.

Lemma 2.7. *Let Λ_N be the Lebesgue constant for the $(N+1)$ Chebyshev points on an interval $[a,b]$. For any nonnegative integer N, the Lebesgue constant Λ_N satisfies*

$$\Lambda_N \le \frac{2}{\pi}\log(N+1)+1. \tag{42}$$

The following lemma provides a sufficient condition for the Chebyshev interpolant of a function to converge geometrically. The proof can be found in, for example, Theorem 8.2 in [32]. Recall that the level set E_ρ for an interval $[a,b]$ is a Bernstein ellipse with parameter ρ, with foci at a and b (see Figure 4a).

Lemma 2.8. *Suppose that $F:[a,b]\to\mathbb{C}$ is analytically continuable to the region E_ρ^o (see Definition 2.1), and satisfies $\|F\|_{L^\infty(E_\rho^o)} \le M$ for some $M\ge 0$. The Nth degree Chebyshev interpolant P_N of F satisfies*

$$\|F-P_N\|_{L^\infty([a,b])} \le \frac{4M}{\rho-1}\rho^{-N}, \tag{43}$$

for all $N\ge 0$.

We note that the lemma above is stronger than Lemma 2.3 when Γ is an interval, as it specifies the constant factor C.

Remark 2.8. The Legendre points exhibit similar characteristics to the Chebyshev points, and can also be effectively utilized for interpolation over an interval.

In the rest of this section, we provide a series of numerical experiments involving interpolation over intervals. In Figure 5, we report the 2-norm of the inverse of the Vandermonde matrices with Chebyshev collocation points, for the domains $\Gamma=[-1,1]$ and $\Gamma=[0,1]$. Note that when $\Gamma=[-1,1]$, we have that $\rho_*=1+\sqrt{2}$ and $\|(V^{(N)})^{-1}\|_2 \le \frac{1}{u}$ for $N\le 43$; when $\Gamma=[0,1]$, we have that $\rho_*=3+2\sqrt{2}$ and $\|(V^{(N)})^{-1}\|_2 \le \frac{1}{u}$ for $N\le 22$. In Figure 6, we interpolate functions which can be resolved by a Chebyshev interpolant of degree $N\le 43$ over $\Gamma=[-1,1]$. In addition to the estimated values of $\|F-P_N\|_{L^\infty([-1,1])}$ and $\|F-\widehat{P}_N\|_{L^\infty([-1,1])}$, we plot two additional curves in each figure: an a priori estimate for stagnation of convergence, i.e., $u\cdot\|F\|_{L^\infty(\partial D_1)}$ (see inequality (41)); and the estimated values of $u\cdot\|a^{(N)}\|_2$ based on inequality (14). In Figure 7, we provide similar experiments for the case where $\Gamma=[0,1]$. In Figure 8, we consider polynomial interpolation, in the monomial basis, of more complicated functions that cannot be resolved to machine precision by a Chebyshev interpolant of degree less than the threshold. Based on these experimental results, we make the following observations:

1. The convergence generally stagnates after the monomial approximation error $\|F-\widehat{P}_N\|_{L^\infty(\Gamma)}$ reaches $u\cdot\|a^{(N)}\|_2$, which implies that inequality (35) is sharp.
2. Stagnation of convergence occurs once the polynomial interpolation error decays to $u\cdot\|F\|_{L^\infty(\partial D_1)}$, which validates the effectiveness of the prediction for stagnation of convergence (41).
3. The monomial basis is generally as effective as a well-conditioned polynomial basis for interpolation, as long as the order does not exceed the threshold. This observation is in line with our analysis in Section 2.1.

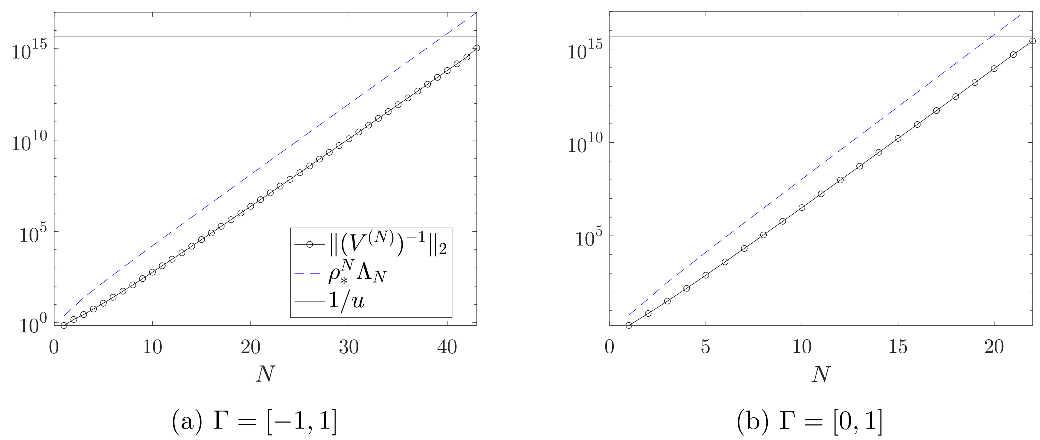
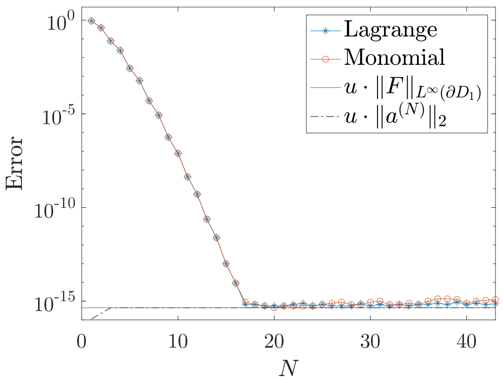


Figure 5: **The 2-norm of the inverse of a Vandermonde matrix with Chebyshev collocation points over an interval Γ, and its upper bound, for different orders of approximation**. We note that $\rho_* = 1 + \sqrt{2}$ when $\Gamma = [-1, 1]$, and $\rho_* = 3 + 2\sqrt{2}$ when $\Gamma = [0, 1]$.

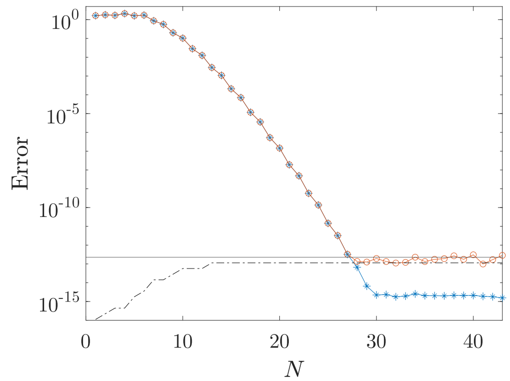
2.4 Interpolation over a smooth simple arc in the complex plane

In this section, we consider polynomial interpolation in the monomial basis over a smooth simple arc $\Gamma \subset \mathbb{C}$. In this more general setting, similar to the special case where Γ is an interval, there exists a class of collocation points, known as adjusted Fejér points, whose associated Lebesgue constant also grows logarithmically [34]. However, these points are extremely costly to construct numerically. On the other hand, the set of collocation points constructed based on the following procedure, while suboptimal, is a good choice for practical applications. Suppose that $g : [-1, 1] \to \mathbb{C}$ is a parameterization of Γ. Provided that the Jacobian $g'(t)$ does not have large variations, we find that the Lebesgue constant for the set of collocation points $Z = \{g(t_j)\}_{j=0,1,\ldots,N}$, where $\{t_j\}_{j=0,1,\ldots,N}$ is the set of $(N+1)$ Chebyshev points on the interval $[-1, 1]$, grows at a slow rate. It is worth noting that $\{t_j\}_{j=0,1,\ldots,N}$ can also be chosen as the Legendre points on the interval $[-1, 1]$, for the same reason stated in Remark 2.8.

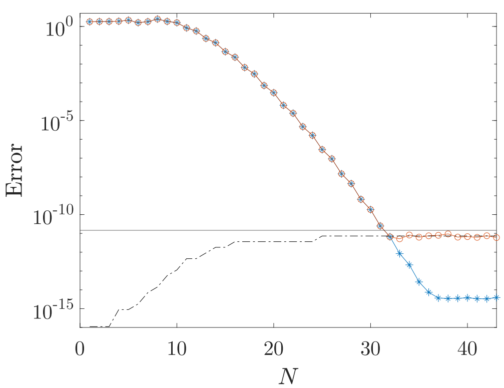
In the rest of this section, we provide several numerical experiments involving interpolation over smooth simple arcs in the complex plane. In particular, we consider the scenario where Γ is a parabola parameterized by $g : [-1, 1] \to \mathbb{C}$, $g(t) := t + i\alpha(t^2 - 1)$, for $\alpha = 0.2$, 0.4, 0.6. In Figure 9, we plot these parabolas, including their associated level sets E_ρ, for various values of ρ. The value of ρ_* for each parabola is estimated from the plots. In Figure 10, we estimate the condition numbers of the Vandermonde matrices and the Lebesgue constants for the sets of collocation points, for different values of α. One can observe that the Lebesgue constants are of approximately size one, which justifies our choice of collocation points. In Figure 11, we report the monomial approximation error $\|F - \widehat{P}_N\|_{L^\infty(\Gamma)}$, an a priori estimate for stagnation of convergence (i.e., $u \cdot \|F\|_{L^\infty(\partial D_1)}$), and the estimated values of $u \cdot \|a^{(N)}\|_2$, for various functions F over Γ. Based on the experimental results, it is clear that the observations made at the end of Section 2.3 are also applicable to the case where Γ is a parabola. In fact, these observations apply to any simple arc that is sufficiently smooth.



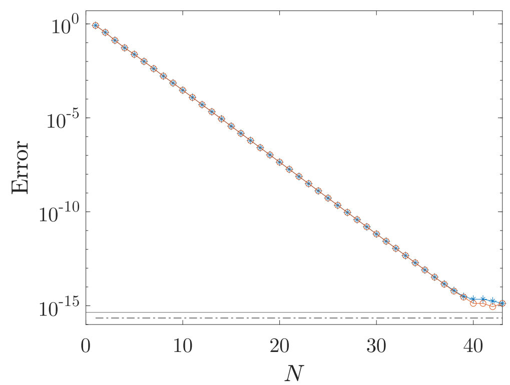
(a) $F(x) = \cos(2x + 1)$



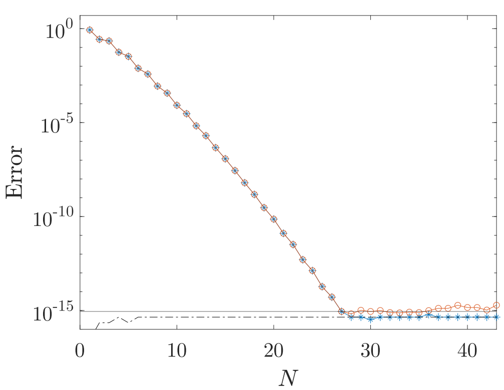
(b) $F(x) = \cos(8x + 1)$



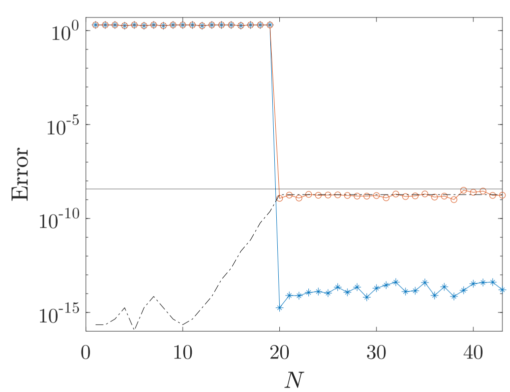
(c) $F(x) = \cos(12x + 1)$



(d) $F(x) = \frac{1}{x - \sqrt{2}}$

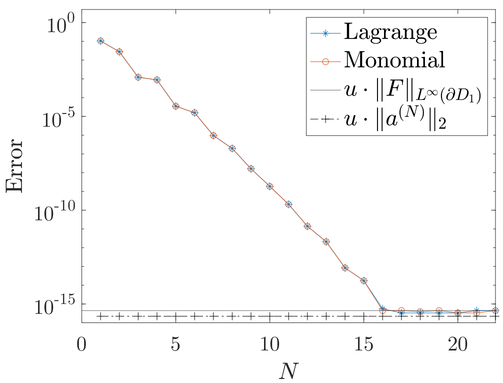


(e) $F(x) = e^{-2(x+0.1)^2}$

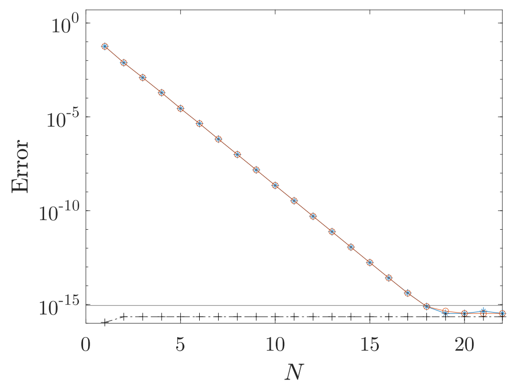


(f) $F(x) = T_{20}(x)$

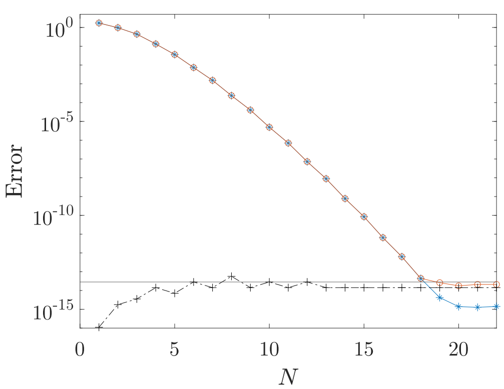
Figure 6: **Polynomial interpolation in the monomial basis over** $\Gamma = [-1, 1]$. The label "Lagrange" denotes $\|F - P_N\|_{L^\infty(\Gamma)}$, estimated using the Barycentric interpolation formula. The label "Monomial" denotes the estimated value of $\|F - \widehat{P}_N\|_{L^\infty(\Gamma)}$. Recall that the functions F in Figures 6a-6d are the same functions that appear in Section 1.



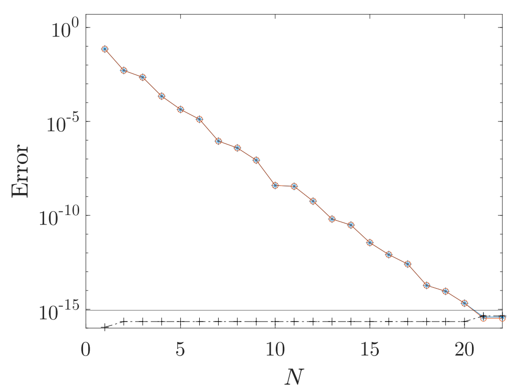
(a) $F(x) = e^{-x^2}$



(b) $F(x) = \frac{1}{x+1.2}$

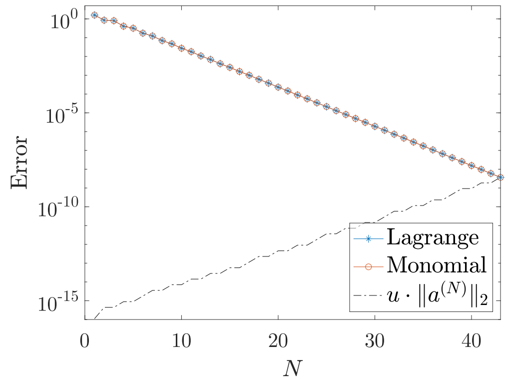


(c) $F(x) = \sin(6x + 1)$

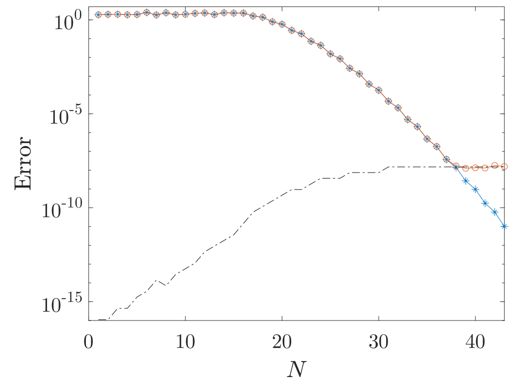


(d) $F(x) = \arctan(x)$

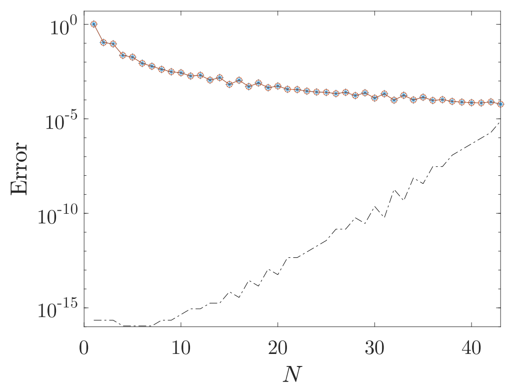
Figure 7: **Polynomial interpolation in the monomial basis over** $\Gamma = [0, 1]$. See the caption of Figure 6 for the definitions of the labels "Lagrange", "Monomial".



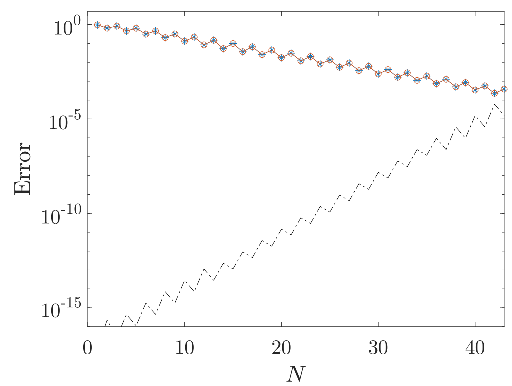
(a) $F(x) = \frac{1}{x-0.5i}$, $\Gamma = [-1, 1]$



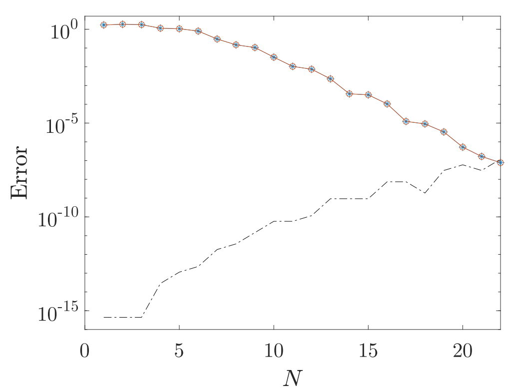
(b) $F(x) = \cos(20x + 1)$, $\Gamma = [-1, 1]$



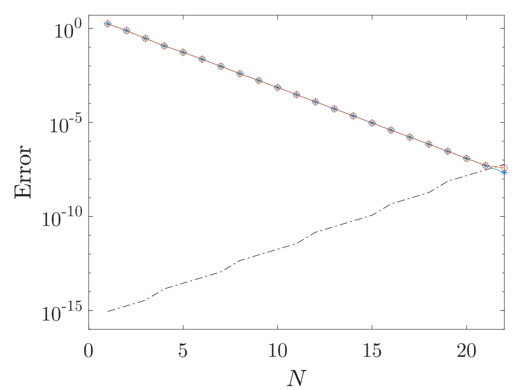
(c) $F(x) = |x + 0.1|^{2.5}$, $\Gamma = [-1, 1]$



(d) $F(x) = \frac{1}{1+25x^2}$, $\Gamma = [-1, 1]$

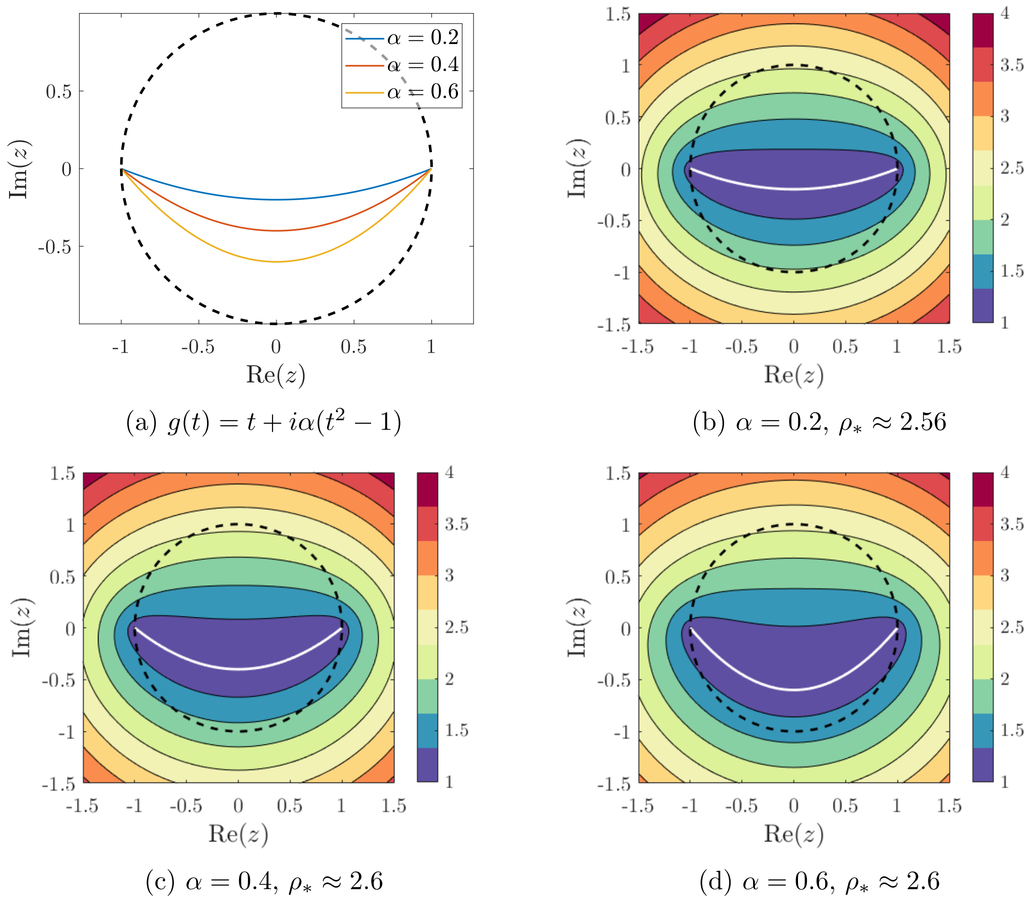


(e) $F(x) = \sin(10x^2 + 1)$, $\Gamma = [0, 1]$



(f) $F(x) = \frac{1}{x+0.2}$, $\Gamma = [0, 1]$

Figure 8: **Polynomial interpolation, in the monomial basis, of functions that are more difficult to resolve**. See the caption of Figure 6 for the definitions of the labels “Lagrange” and “Monomial”.



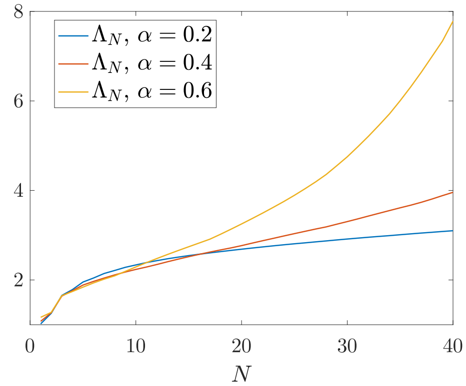
(a) $g(t) = t + i\alpha(t^2 - 1)$

(b) $\alpha = 0.2$, $\rho_* \approx 2.56$

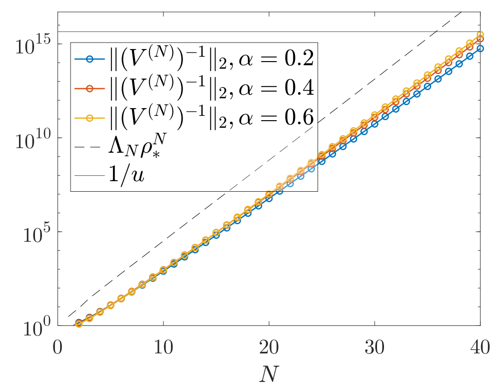
(c) $\alpha = 0.4$, $\rho_* \approx 2.6$

(d) $\alpha = 0.6$, $\rho_* \approx 2.6$

Figure 9: **The level set E_ρ of a parabola, for various values of ρ.** The colorbar indicates the value of ρ. The smooth simple arc Γ is the white curve in the figure. The value of ρ_* (see Definition 2.2) is estimated for each arc Γ.



(a)



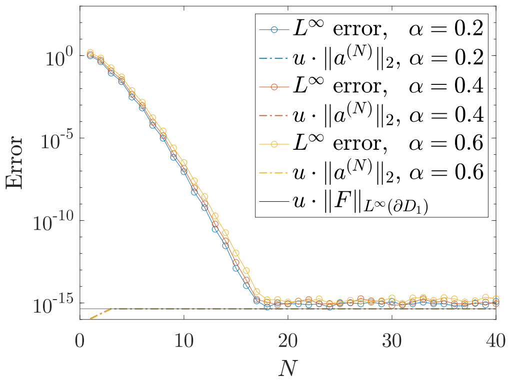
(b)

Figure 10: **The Lebesgue constant for collocation points over a parabola, and the 2-norm of the inverse of the corresponding Vandermonde matrix**. The collocation points are chosen to be $\{g(t_j)\}$, where $\{t_j\}$ is a set of Chebyshev points over $[-1, 1]$, and $g : [-1, 1] \to \mathbb{C}$ is the parameterization of the parabola defined in Section 2.4. The x-axis label N denotes the order of approximation. The value of ρ_* is set to be 2.6, based on the estimate in Figure 9.

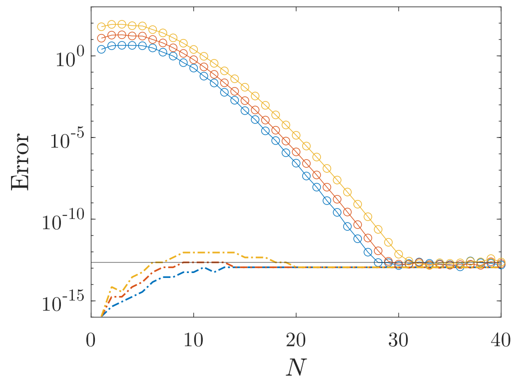
Remark 2.9. In certain applications, the function $F : \Gamma \to \mathbb{C}$ is defined by the formula $F(z) := \sigma(g^{-1}(z))$, where $g : [-1, 1] \to \mathbb{C}$ is an analytic function that parameterizes the curve Γ, and $\sigma : [-1, 1] \to \mathbb{C}$ is analytic. In this case, the analytic continuation of F can have a singularity close to Γ even when σ is entire, because the inverse of the parameterization (i.e., g^{-1}) has so-called Schwarz singularities at $z = g(t^*)$, where $g'(t^*) = 0$. In [2], the authors show that, the higher the curvature of the arc Γ, the closer the singularity induced by g^{-1} is to Γ. As a result, the approximation of such a function F by polynomials is efficient only when the curvature of Γ is small.

3 Applications

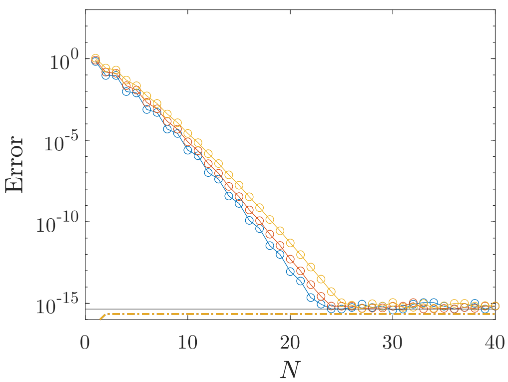
After justifying the use of a monomial basis for polynomial interpolation, a natural question to ask is: why would one want to do it in the first place? For one, the monomial basis is the simplest polynomial basis to manipulate. For example, the evaluation of an Nth degree polynomial expressed in the monomial basis can be achieved using only N multiplications through the application of Horner's rule. Additionally, the derivative and anti-derivative of an Nth degree polynomial in the monomial basis can be calculated more stably in other bases, and using only N multiplications. Besides these obvious advantages, we present several applications that demonstrate the unique merits of polynomial interpolation in the monomial basis.



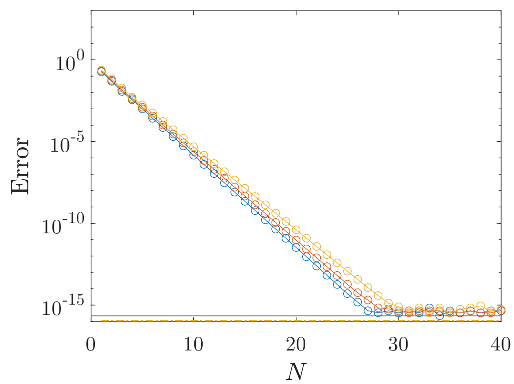
(a) $F(z) = \cos(2z + 1)$



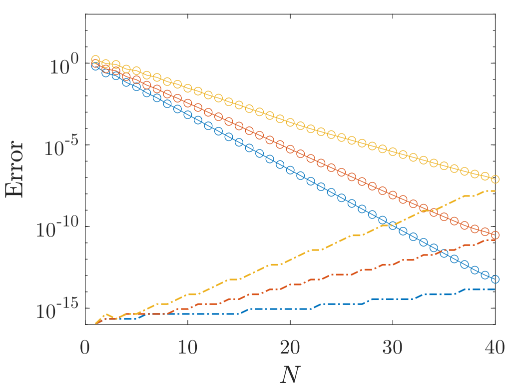
(b) $F(z) = \cos(8z + 1)$



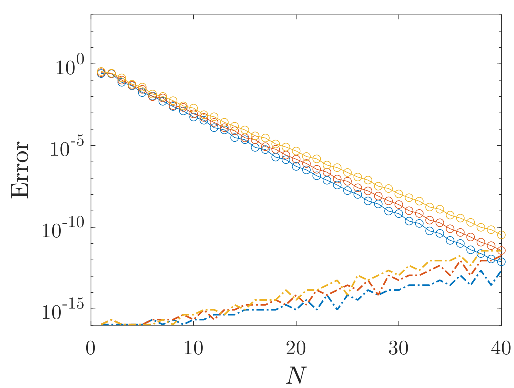
(c) $F(z) = e^{-z^2}$



(d) $F(z) = \frac{1}{z-2}$



(e) $F(z) = \frac{1}{z+i}$



(f) $F(z) = \tan(\tan(z)/2)$

Figure 11: **Polynomial interpolation in the monomial basis over a parabola**. The interpolation is performed on the parabolas shown in Figure 9a. The x-axis label N denotes the order of approximation. The label "L^∞ error" denotes the estimated value of $\|F - \widehat{P}_N\|_{L^\infty(\Gamma)}$.

3.1 Oscillatory integrals and singular integrals

Given an oscillatory (or singular) function $\Psi : \Gamma \to \mathbb{C}$ and a smooth function $F : \Gamma \to \mathbb{C}$ over a smooth simple arc $\Gamma \subset \mathbb{C}$, the calculation of

$$\int_\Gamma \Psi(z)F(z)\,\mathrm{d}z \tag{44}$$

by standard quadrature rules can be extremely expensive or inaccurate due to the oscillations (or the singularity) of Ψ. However, when F is a monomial, there exists a wide range of integrals in the form (44) that can be efficiently computed to high accuracy by either analytical formulas or by recurrence relations, often derived using integration by parts. Therefore, when the smooth function F is accurately approximated by a monomial expansion of order N, such integrals can be efficiently evaluated by the formula

$$\sum_{k=0}^{N} a_k \Big(\int_\Gamma \Psi(z) z^k \,\mathrm{d}z \Big), \tag{45}$$

where $\{a_k\}_{k=0,1,\ldots,N}$ denotes the coefficients of the monomial expansion.

In the rest of this section, we present examples of oscillatory integrals and singular integrals of this kind.

Remark 3.1. When one needs to compute (44) for multiple smooth functions F, it is unnecessary to compute a monomial expansion for each F. Instead, the adjoint method can be used to compute a quadrature rule $\{(z_i, w_i)\}_{i=0,1,\ldots,N}$, such that

$$\int_\Gamma \Psi(z)F(z)\,\mathrm{d}z \approx \sum_{i=0}^{N} w_i \Psi(z_i)F(z_i), \tag{46}$$

for any function F that can be accurately approximated by a monomial expansion of order N. We refer the readers to Section 2.2.2 in [2] for a detailed overview of the method.

3.1.1 Fourier integrals

Given a smooth function $G : [a, b] \to \mathbb{C}$ and a real number c, it takes $\mathcal{O}(c)$ operations to compute the Fourier integral

$$\int_a^b e^{icx} G(x)\,\mathrm{d}x \tag{47}$$

by a standard quadrature rule for smooth functions, as the number of points required to resolve the integrand is proportional to the value of c. Consequently, the evaluation of such an integral is prohibitively expensive when c is large. By a change of variables, the integral (47) can be decomposed into the sum of the integral of a smooth function, and an oscillatory integral of the form

$$\int_{-1}^1 e^{i\omega x} F(x)\,\mathrm{d}x, \tag{48}$$

where $\omega \in \mathbb{R}$, and $F : [-1, 1] \to \mathbb{R}$ is smooth. Thus, without loss of generality, it is sufficient to consider the numerical evaluation of (48) alone. It is known that (48) can be efficiently evaluated to high accuracy with a cost independent of ω when the function F is a monomial, using the following recurrence relations:

$$\int_{-1}^{1} e^{i\omega x}\,\mathrm{d}x = \frac{1}{i\omega}(e^{i\omega} - e^{-i\omega}), \tag{49a}$$

$$\int_{-1}^{1} e^{i\omega x}x^{k+1}\,\mathrm{d}x = \frac{1}{i\omega}\Big(e^{i\omega} + (-1)^k e^{-i\omega} - (k+1)\int_{-1}^{1} e^{i\omega x}x^k\,\mathrm{d}x\Big), \tag{49b}$$

for all $k \geq 0$. The use of this recurrence relation for computing Fourier integrals was first proposed in [14], and the resulting algorithm is known as the Filon-type method. When this method is used, the smooth function F is typically approximated by piecewise polynomials of low degrees (typically less than five) [14, 23], in part due to the belief that higher-order polynomial interpolation in the monomial basis is unstable. By approximating F by a higher-order monomial expansion as described in this paper, the Filon-type method is made substantially more accurate.

We note that this technique also generalizes to higher dimensions and to more complicated oscillatory functions Ψ, and we refer the readers to [24] for an overview.

3.1.2 Layer potentials

Given a target point $\xi \in \mathbb{C}$, and a smooth simple arc $\Gamma \subset \mathbb{C}$ with endpoints z_1 and z_2, the evaluation of the layer potentials

$$\int_\Gamma \log(z-\xi)F(z)\,\mathrm{d}z \quad \text{and} \quad \int_\Gamma \frac{F(z)}{z-\xi}\,\mathrm{d}z, \tag{50}$$

is of great importance in the integral equation method for the numerical solution of partial differential equations [25]. Without loss of generality, we assume that $z_1 = -1$ and $z_2 = 1$. When the target ξ is close to Γ, the integrands of (50) become nearly singular and, as a result, standard quadrature rules cannot be used to compute the integrals efficiently. In [20, 18], the authors observe that, when the function F is a monomial, the integrals (50) satisfy the following recurrence relations:

$$\int_\Gamma \frac{1}{z-\xi}\,\mathrm{d}z = \log(1-\xi) - \log(-1-\xi) + 2\pi i\mathcal{N}_\xi, \tag{51a}$$

$$\int_\Gamma \frac{z^{k+1}}{z-\xi}\,\mathrm{d}z = \xi\int_\Gamma \frac{z^k}{z-\xi}\,\mathrm{d}z + \frac{1+(-1)^k}{k+1}, \tag{51b}$$

$$\int_\Gamma \log(z-\xi)z^k\,\mathrm{d}z = \frac{1}{k+1}\Big(\log(1-\xi) + (-1)^k\log(-1-\xi) - \int_\Gamma \frac{z^{k+1}}{z-\xi}\,\mathrm{d}z\Big), \tag{51c}$$

for all $k \geq 0$, where $\mathcal{N}_\xi \in \mathbb{Z}$ is a winding number determined by the position of ξ relative to Γ. Moreover, these recurrence relations are stable when ξ is close to Γ. Consequently, the layer potentials (50) can be efficiently evaluated by interpolating F in the monomial basis. We refer the readers to Section 2.2.1 in [2] for a comprehensive overview of this method. It is worth noting that other types of layer potentials can be computed using a similar approach, as discussed in [19, 28].

3.1.3 Hadamard finite-part integrals

Integrals of the form

$$\int_a^b (x-a)^\nu \log^m(x-a) G(x)\,\mathrm{d}x, \tag{52}$$

where $G\colon [a,b] \to \mathbb{C}$ is smooth, $\nu \in \mathbb{R}$, and $m \geq 0$ is an integer, appear in numerous applications. By a change of variables, the integral (52) can be written as a combination of integrals of the form

$$\int_0^1 x^\nu \log^m(x) F(x)\,\mathrm{d}x, \tag{53}$$

where $F\colon [0,1] \to \mathbb{C}$ is smooth. When $\nu \leq -1$, this integral is divergent, in which case we can consider only its "finite part" (see, for example, [13]). Let $\epsilon > 0$, and write

$$\int_\epsilon^1 x^\nu \log^m(x) F(x)\,\mathrm{d}x = F_0(\epsilon) + F_1(\epsilon), \tag{54}$$

where $F_0(\epsilon)$ remains bounded as $\epsilon \to 0$, and

$$F_1(\epsilon) = a_1\Psi_1(\epsilon) + a_2\Psi_2(\epsilon) + \cdots + a_n\Psi_n(\epsilon) \tag{55}$$

is a combination of given functions $\Psi_1, \Psi_2, \ldots, \Psi_n$ which become infinite as $\epsilon \to 0$. Discarding the "infinite part" $F_1(\epsilon)$, we define the Hadamard finite part of (53) by

$$\text{f.p.} \int_0^1 x^\nu \log^m(x) F(x)\,\mathrm{d}x = \lim_{\epsilon \to 0} F_0(\epsilon). \tag{56}$$

It is possible to show that the finite part of (53) is equal to its meromorphic continuation in ν to the region $\{\nu \in \mathbb{C} : \nu \neq -1, -2, \ldots\}$.

When $F(x)$ is a monomial, we can evaluate the finite part explicitly using the formula

$$\text{f.p.} \int_0^1 x^\nu \log^m(x) \cdot x^k\,\mathrm{d}x = \frac{(-1)^m m!}{(\nu+k+1)^{m+1}}. \tag{57}$$

Therefore, the finite part integral (53) can be accurately and efficiently evaluated for all $\nu \notin \{-1, -2, \ldots\}$, once $F(x)$ is approximated by monomials.

3.2 Root finding

Given a smooth simple arc $\Gamma \subset \mathbb{C}$ and a function $F : \Gamma \to \mathbb{C}$, one method for computing the roots of F over Γ is to first approximate it by a polynomial $P_N(z) = \sum_{j=0}^N a_j z^j$ to high accuracy, and then to compute the roots of P_N by calculating the eigenvalues of the companion matrix

$$C(P_N) := \begin{pmatrix} 0 & 0 & \ldots & 0 & -\frac{a_0}{a_N} \\ 1 & 0 & \ldots & 0 & -\frac{a_1}{a_N} \\ 0 & 1 & \ldots & 0 & -\frac{a_2}{a_N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -\frac{a_{N-1}}{a_N} \end{pmatrix}. \tag{58}$$

Recently, a backward stable algorithm that computes the eigenvalues of $C(P_N)$ in $\mathcal{O}(N^2)$ operations with $\mathcal{O}(N)$ storage has been proposed in [5]. This algorithm is backward stable in the sense that the computed roots are the exact roots of a perturbed polynomial $\widehat{P}_N(z) = \sum_{j=0}^{N}(a_j + \delta a_j)z^j$, so that the backward error satisfies $\|\delta a^{(N)}\|_2 \lesssim u\|a^{(N)}\|_2$, where u denotes machine epsilon, $\delta a^{(N)} := (\delta a_0, \delta a_1, \ldots, \delta a_N)^T$ and $a^{(N)} := (a_0, a_1, \ldots, a_N)^T$. It follows that

$$\|P_N - \widehat{P}_N\|_{L^\infty(\Gamma)} \leq u\|\delta a^{(N)}\|_1 \lesssim u\sqrt{N+1}\|a^{(N)}\|_2. \tag{59}$$

When $\|a^{(N)}\|_2 \approx \|P_N\|_{L^\infty(\Gamma)}$, the computed roots are backward stable in the polynomial P_N. This condition, however, does not hold for all polynomials P_N. Furthermore, the calculation of the coefficients $a^{(N)}$ from the function F, which involves the solution of a Vandermonde system of equations, is highly ill-conditioned.

In this paper, we show that, when F is sufficiently smooth, it is possible to compute the coefficients of an interpolating polynomial $P_N(z) = \sum_{j=0}^{N} a_j z^j$, with $\|a^{(N)}\|_2 \approx \|F\|_{L^\infty(\Gamma)}$, which approximates F uniformly to high accuracy, even when N is moderately large. From this, we see that a backward stable root finder can be constructed by combining the piecewise polynomial approximation procedure described in Section 2.2 with the algorithm presented in [5],

Remark 3.2. We note that, in the case when Γ is an interval, the Chebyshev polynomial basis is more advantageous than the monomial basis for interpolation in the context of root finding, since the relationship between the values of a polynomial and its Chebyshev coefficients is stable, and there exist backward stable algorithms which likewise compute the roots of an Nth degree polynomial in the Chebyshev polynomial basis in $\mathcal{O}(N^2)$ operations with $\mathcal{O}(N)$ storage [29].

4 Discussion

Since the invention of digital computers, most research on the topic of polynomial interpolation in the monomial basis focuses on showing that it is a bad idea. The condition number of Vandermonde matrices has been studied extensively in recent decades (see [15] for a literature review), and it is known that its growth rate is at least exponential, unless the collocation nodes are distributed uniformly on the unit circle centered at the origin [27]. As a result, the computed monomial coefficients are generally highly inaccurate when the dimensionality of the Vandermonde matrix is not small. For this reason, other more well-conditioned bases are often used for polynomial interpolation [32, 11]. However, the fact that the monomial coefficients are computed inaccurately does not imply that polynomial interpolation in the monomial basis is unstable, since it is the backward error $\|V\widehat{a} - f\|_2$ of the numerical solution $\widehat{a}$ to the Vandermonde system $Va = f$ that determines the accuracy of the approximation, and $\|V\widehat{a} - f\|_2$ can be small even when the condition number $\kappa(V)$ is large (similar situations also occur in the method of fundamental solutions [6, 31] and in frame approximations [3, 4]). It has long been observed that, as a result, polynomial interpolation in the monomial basis produces highly accurate approximations for sufficiently smooth functions (see, for example, [17, 20]), and many interesting applications have appeared, for example, [21, 28, 1, 26]. Yet, the general attitude towards it has remained skeptical, in part because a complete theory has been unavailable until now. We show in this paper that, so long as $\kappa(V) \lesssim \frac{1}{u}$ and a backward stable linear system solver is used, the approximation error of the computed monomial expansion is bounded by approximately the sum of the exact polynomial interpolation error and an extra error term $u\|a\|_2$. Our key observation is that this extra error term generally does not cause the monomial basis to be inferior to a well-conditioned basis for interpolation, given that the order is no larger than the maximum order allowed by the constraint $\kappa(V) \lesssim \frac{1}{u}$. Since this maximum order is not small in practice, we find that the monomial basis is a useful basis for interpolation, especially when it is used to construct a piecewise polynomial approximation. Besides showing that the monomial basis can be used stably, we present a number of applications where the use of a monomial basis for interpolation offers a substantial advantage over other bases.

While not discussed in this paper, our theory can be easily generalized to higher dimensions if an analogue of Lemma 2.4 is available. We conjecture that such an analogue exists, as we observe that, similar to the univariate case, the monomial basis is generally as good as a well-conditioned polynomial basis for multivariate interpolation, provided that the order is below a certain threshold analogous to the one described in the univariate case. Our recent work on Newtonian potential evaluation [30] relies heavily on this observation, where the monomial basis is used for the approximation of the anti-Laplacian of a 2-D function.

5 Acknowledgements

We are deeply grateful to James Bremer, Andreas Klöeckner, Nick Trefethen, and Daan Huybrechs for their valuable advice and insightful discussions.

References

[1] af Klinteberg, L., Askham, T., Kropinski, M. C.: A fast integral equation method for the two-dimensional Navier-Stokes equations. J. Comput. Phys., **409**, 109353 (2020)

[2] af Klinteberg, L., Barnett, A. H.: Accurate Quadrature of Nearly Singular Line Integrals in Two and Three Dimensions by Singularity Swapping. BIT Numer. Math. **61**.1, 83–118 (2021)

[3] Adcock, B., Huybrechs, D.: Frames and Numerical Approximation. SIAM Rev. **61**(3), 443–473 (2019)

[4] Adcock, B., Huybrechs, D.: Frames and Numerical Approximation II: Generalized Sampling. J. Fourier Anal. Appl. **26**(6), 1–34 (2020)

[5] Aurentz, J. L., Mach, T., Vandebril, R., Watkins, D. S.: Fast and Backward Stable Computation of Roots of Polynomials. SIAM J. Matrix Anal. Appl. **36**(3), 942–973 (2015)

[6] Barnett, A. H., Betcke, T.: Stability and Convergence of the Method of Fundamental Solutions for Helmholtz Problems on Analytic Domains. J. Comput. Phys. **227**(14), 7003–7026 (2008)

[7] Beckermann, B.: The Condition Number of Real Vandermonde, Krylov and Positive Definite Hankel Matrices. Numer. Math. **85**(4), 553–577 (2000)

[8] Berrut, J.-P., Trefethen, L. N.: Barycentric Lagrange Interpolation. SIAM Rev. **46**(3), 501–517 (2004)

[9] Björck, Å., Pereyra, V.: Solution of Vandermonde Systems of Equations. Math. Comput. **24**(112), 893–903 (1970)

[10] Börm, S.: On Iterated Interpolation. SIAM J. Numer. Anal. **60**(6), 3124–3144 (2022)

[11] Brubeck, P. D., Nakatsukasa Y., Trefethen, L. N.: Vandermonde with Arnoldi, SIAM Rev. **63**(2), 405–415 (2021)

[12] Ehlich, H., Zeller, K.: Auswertung der Normen von Interpolationsoperatoren. Math. Ann. **164**, 105–112 (1966)

[13] Estrada, R., Kanwal, R. P.: Singular Integral Equations. Springer Science & Business Media (2012)

[14] Filon, L. N. G.: On a Quadrature Formula for Trigonometric Integrals. Proc. R. Soc. Edinb. **49**, 38–37 (1930)

[15] Gautschi, W.: How (un)stable are Vandermonde systems?. Asymptot. Comput. Anal. **124**, 193–210 (1990)

[16] Gohberg, I., Olshevsky, V.: The Fast Generalized ParkerTraub Algorithm for Inversion of Vandermonde and Related Matrices. J. Complex. **13**(2), 208–234 (1997)

[17] Heath, M. T.: Scientific computing: an introductory survey, revised second edition. SIAM (2018)

[18] Helsing, J.: Integral Equation Methods for Elliptic Problems with Boundary Conditions of Mixed Type. J. Comput. Phys. **228**(23), 8892–8907 (2009)

[19] Helsing, J., Holst, A.: Variants of an Explicit Kernel-Split Panel-Based Nystrm Discretization Scheme for Helmholtz Boundary Value Problems. Adv. Comput. Mathe. **41**(3), 691–708 (2015)

[20] Helsing, J., Ojala, R.: On the Evaluation of Layer Potentials Close to Their Sources. J. Comput. Phys. **227**(5), 2899–2921 (2008)

[21] Helsing, J., Jiang, S.: On Integral Equation Methods for the First Dirichlet Problem of the Biharmonic and Modified Biharmonic Equations in NonSmooth Domains. SIAM J. Sci. Comput. **40**(4), A2609–2630 (2018)

[22] Higham, N. J.: The Numerical Stability of Barycentric Lagrange Interpolation. IMA J. Numer. Anal. **24**(4), 547–556 (2004)

[23] Iserles, A., Nørsett, S. P.: Efficient Quadrature of Highly Oscillatory Integrals Using Derivatives. Proc. Math. Phys. Eng. **461**(2057), 1383–1399 (2005)

[24] Iserles, A., Nørsett, S. P., Olver, S.: Highly Oscillatory Quadrature: The Story So Far. In: Proceedings of ENumath, Santiago de Compostela (2006). Springer, Berlin, pp. 97–118 (2006)

[25] Martinsson, P. G.: Fast Direct Solvers for Elliptic PDEs. CBMS-NSF Conference Series, vol. CB96. SIAM, Philadelphia (2019)

[26] Ojala, R., Tornberg, A.-K.: An Accurate Integral Equation Method for Simulating Multi-Phase Stokes Flow. J. Comp. Phys. **298**, 145–160 (2015)

[27] Pan, V. Y.: How bad are Vandermonde matrices?, SIAM J. Matrix Anal. Appl. **37**(2), 676–694 (2016)

[28] Wu, B., Zhu, H., Barnett, A., Veerapaneni S.: Solution of Stokes Flow in Complex Nonsmooth 2D Geometries via a Linear-Scaling High-Order Adaptive Integral Equation Scheme. J. Comput. Phys. **410**, pp. 109361 (2020)

[29] Serkh, K., Rokhlin, V.: A Provably Componentwise Backward Stable $\mathcal{O}(n^2)$ QR Algorithm for the Diagonalization of Colleague Matrices. arXiv:2102.12186 (2021)

[30] Shen, Z, Serkh, K.: Rapid evaluation of Newtonian potentials on planar domains. arXiv:2208.10443 (2022)

[31] Stein, D. B., Barnett, A. H.: Quadrature by Fundamental Solutions: Kernel-Independent Layer Potential Evaluation for Large Collections of Simple Objects. Adv. Comput. Math. **48**(60) (2022)

[32] Trefethen, L. N.: Approximation Theory and Approximation Practice. SIAM (2019)

[33] Walsh, J. L.: Interpolation and Approximation by Rational Functions in the Complex Domain. vol. 20. AMS, Philadelphia (1935)

[34] Zhong, L. F., Zhu, L. Y.: The Marcinkiewicz-Zygmund Inequality on a Smooth Simple Arc. J. Approx. Theory **83**(1), 65–83 (1995)