Found Comput Math (2014) 14:635–687
DOI 10.1007/s10208-013-9158-8

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

On the Numerical Stability of Fourier Extensions

**Ben Adcock · Daan Huybrechs ·
Jesús Martín-Vaquero**

Received: 13 June 2012 / Revised: 20 March 2013 / Accepted: 2 May 2013 /
Published online: 15 April 2014

Abstract An effective means to approximate an analytic, nonperiodic function on a bounded interval is by using a Fourier series on a larger domain. When constructed appropriately, this so-called Fourier extension is known to converge geometrically fast in the truncation parameter. Unfortunately, computing a Fourier extension requires solving an ill-conditioned linear system, and hence one might expect such rapid convergence to be destroyed when carrying out computations in finite precision. The purpose of this paper is to show that this is not the case. Specifically, we show that Fourier extensions are actually numerically stable when implemented in finite arithmetic, and achieve a convergence rate that is at least superalgebraic. Thus, in this instance, ill-conditioning of the linear system does not prohibit a good approximation.

In the second part of this paper we consider the issue of computing Fourier extensions from equispaced data. A result of Platte et al. (SIAM Rev. 53(2):308–318, 2011) states that no method for this problem can be both numerically stable and exponentially convergent. We explain how Fourier extensions relate to this theoretical barrier,

Communicated by Nira Dyn.

B. Adcock (✉)
Department of Mathematics, Purdue University, West Lafayette, IN, USA
e-mail: adcock@purdue.edu

D. Huybrechs
Department of Computer Science, Katholieke Universiteit Leuven, Leuven, Belgium

J. Martín-Vaquero
Department of Applied Mathematics, E.T.S.I.I. Béjar, University of Salamanca, Salamanca, Spain

and demonstrate that they are particularly well suited for this problem: namely, they obtain at least superalgebraic convergence in a numerically stable manner.

Keywords Fourier series · Fourier extension · Convergence · Stability · Equispaced data

Mathematics Subject Classification (2010) 42A10 · 65T40 · 42C15

Symbols

T	Extension parameter
N	Truncation parameter
M, γ	Number of equispaced nodes of the equispaced FE, and the oversampling parameter $\gamma = M/N$
$\phi_n(x)$	The exponential $\frac{1}{\sqrt{2T}} \mathrm{e}^{\mathrm{i}\frac{n\pi}{T}x}$
$\mathcal{G}_N, \mathcal{S}_N, \mathcal{C}_N$	Finite-dimensional spaces of exponentials, sines and cosines
F_N, $\tilde{F}_N(f)$, $F_{N,M}(f)$	Exact continuous, discrete and equispaced FEs
G_N, $\tilde{G}_N(f)$, $G_{N,M}(f)$	Numerical continuous, discrete and equispaced FEs
a	Vector of coefficients of an FE
A, $\tilde{A}$, $\bar{A}$	Matrices of the continuous, discrete and equispaced FE's
b, $\tilde{b}$, $\bar{b}$	Data vectors for the continuous, discrete and equispaced FEs
x, y, z	Physical domain variable $x \in [-1, 1]$, and the mapped variables $y \in [c(T), 1]$ and $z \in [-1, 1]$
$f_{\mathrm{e}}(x)$, $f_{\mathrm{o}}(x)$	Even and odd parts of the function $f(x)$
$g_1(y)$, $g_2(y)$, $g_{1,N}(y)$, $g_{2,N}(y)$	Images of $f_{\mathrm{e}}(x)$ and $f_{\mathrm{o}}(x)/\sin\frac{\pi}{T}x$ in the y-domain and their polynomial approximations
$h_i(z)$, $h_{i,N}(z)$	Images of g_i and $g_{i,N}$ in the z-domain
$m(x)$	The mapping $x \mapsto z$
$c(T)$, $E(T)$	FE constants $\cos\frac{\pi}{T}$ and $\cot^2(\frac{\pi}{4T})$.
$\mathcal{B}(\rho)$, $\mathcal{D}(\rho)$	Bernstein ellipse in the z-domain and its image in the x-domain
$\kappa(F)$	Condition number of a mapping F
N_0, N_1, N_2	Breakpoints in convergence
$\{u_n, \sigma_n, v_n\}$	Singular system of A, $\tilde{A}$ or $\bar{A}$
Φ_n	Fourier series corresponding to v_n
$\mathcal{G}_{N,\epsilon}$, $\mathcal{G}'_{N,\epsilon}$, $\mathcal{G}_{N,M,\epsilon}$	The subspace $\mathrm{span}\{\Phi_n : \sigma_n > \epsilon\}$
$H_{N,\epsilon}(f)$, $\tilde{H}_{N,\epsilon}(f)$, $H_{N,M,\epsilon}(f)$	Truncated SVD FEs corresponding to the continuous, discrete and equispaced cases
$a(\gamma; T)$	Quantity determining the maximal achievable accuracy of the equispaced FE
$\mathrm{L}^2(I)$, $\langle\cdot,\cdot\rangle_I$, $\|\cdot\|_I$	Space of square-integral functions on a domain I and corresponding inner product and norm
$\langle\cdot,\cdot\rangle$, $\|\cdot\|$	Inner product and norm on $\mathrm{L}^2(-1, 1)$

$\mathrm{L}^2_w(I)$, $\langle\cdot,\cdot\rangle_{w,I}$, $\|\cdot\|_{w,I}$	Space of square integrable functions with respect to a weight function w and corresponding inner product and norm
$\|\cdot\|_{\infty,I}$, $\|\cdot\|_\infty$	Uniform norms on an arbitrary domain I and the interval $[-1,1]$ respectively

1 Introduction

Let $f:[-1,1]\to\mathbb{R}$ be an analytic function. When periodic, an extremely effective means to approximate f is via its truncated Fourier series. This approximation converges geometrically fast in the truncation parameter N, and can be computed efficiently via the Fast Fourier Transform (FFT). Moreover, Fourier series possess high resolution power. One requires an optimal two modes per wavelength to resolve oscillations, making Fourier methods well suited for (most notably) PDEs with oscillatory solutions [19].

For these reasons, Fourier series are extremely widely used in practice. However, the situation changes completely when f is nonperiodic. In this case, rather than geometric convergence, one witnesses the familiar Gibbs phenomenon near $x=\pm 1$ and only linear pointwise convergence in $(-1,1)$.

1.1 Fourier Extensions

For analytic and nonperiodic functions, one way to restore the good properties of a Fourier series expansion (in particular, geometric convergence and high resolution power) is to approximate f with a Fourier series on an *extended* domain $[-T,T]$. Here $T>1$ is a user-determined parameter. Thus we seek an approximation $F_N(f)$ to f from the set

$$\mathcal{G}_N := \operatorname{span}\bigl\{\phi_n : |n|\le N\bigr\}, \quad \phi_n(x) := \frac{1}{\sqrt{2T}}\mathrm{e}^{\mathrm{i}\frac{n\pi}{T}x}.$$

Although there are many potential ways to define $F_N(f)$, in [5, 12, 22] it was proposed to compute $F_N(f)$ as the best approximation to f on $[-1,1]$ in a least squares sense:

$$F_N(f) := \underset{\phi\in\mathcal{G}_N}{\operatorname{argmin}}\|f-\phi\|. \tag{1.1}$$

Here $\|\cdot\|$ is the standard norm on $\mathrm{L}^2(-1,1)$—the space of square-integrable functions on $[-1,1]$. Henceforth, we shall refer to $F_N(f)$ as the *continuous* Fourier extension (FE) of f.

In [1, 22] it was shown that the continuous FE $F_N(f)$ converges geometrically fast in N and has a resolution constant (number of degrees of freedom per wavelength required to resolve an oscillatory wave) that ranges between 2 and π depending on the choice of the parameter T, with $T\approx 1$ giving close to the optimal value 2 (see Sect. 2.4 for a discussion). Thus the continuous FE successfully retains the key properties of rapid convergence and high resolution power of a standard Fourier series in the case of nonperiodic functions.

We note that one does not usually compute the continuous FE (1.1) in practice. A more convenient approach [1, 22] is to replace (1.1) by the discrete least squares

$$\tilde{F}_N(f) := \underset{\phi\in\mathcal{G}_N}{\operatorname{argmin}} \sum_{|n|\leq N} \left|f(x_n)-\phi(x_n)\right|^2, \tag{1.2}$$

for nodes $\{x_n\}_{|n|\leq N} \subseteq [-1,1]$. We refer to $\tilde{F}_N(f)$ as the *discrete* Fourier extension of f. When chosen suitably—in particular, as in (2.11)—such nodes ensure that the difference in approximation properties between the extensions (1.1) and (1.2) is minimal (for details, see Sect. 2.2).

1.2 Numerical Convergence and Stability of Fourier Extensions

The approximation properties of the continuous and discrete FEs were analyzed in [1, 22]. Therein it was also observed numerically that the condition numbers of the matrices A and $\tilde{A}$ of the least squares (1.1) and (1.2) are exponentially large in N. We shall confirm this observation later in the paper. Thus, if $a=(a_{-N},\ldots,a_N)^\top$ is the vector of *coefficients* of the continuous or discrete FE, i.e. $F_N(f)$ or $\tilde{F}_N(f)$ is given by $\sum_{|n|\leq N} a_n\phi_n$, one expects small perturbations in f to lead to large errors in a. In other words, the computation of the coefficients of the continuous or discrete FE is ill-conditioned.

Because of this ill-conditioning, it is tempting to think that FEs will be useless in applications. At first sight it is reasonable to expect that the good approximation properties of *exact* FEs (i.e. those obtained in exact arithmetic) will be destroyed when computing *numerical* FEs in finite precision. However, previous numerical studies [1, 5, 12, 22, 24, 25] indicate otherwise. Despite very large condition numbers, one typically obtains an extremely good approximation with a numerical FE, even for poorly behaved functions and in the presence of noise.

The aim of this paper is to give a full explanation of this phenomenon. This explanation can be summarized as follows. In computations, one's interest does not lie with the accuracy in computing the coefficient vector a, but rather the accuracy of the numerical FE approximation $\sum_{|n|\leq N} a_n\phi_n$. As we show, although the mapping from a function to its coefficients is ill-conditioned, the mapping from f to its numerical FE is, in fact, well-conditioned. In other words, whilst the small singular values of A (or $\tilde{A}$) have a substantial effect on a, they have a much less significant, and completely quantifiable, effect on the FE itself.

Although this observation explains the apparent stability of numerical FEs, it does not address their approximation properties. In [1, 22] it was shown that the exact continuous and discrete FEs $F_N(f)$ and $\tilde{F}_N(f)$ converge geometrically fast in N. However, the fact that there may be substantial differences between the coefficients of $F_N(f)$, $\tilde{F}_N(f)$ and those of the numerical FEs, which henceforth we denote by $G_N(f)$ and $\tilde{G}_N(f)$, suggests that geometric convergence may not be witnessed in finite arithmetic for large N. As we show later, for a large class of functions, geometric convergence of $F_N(f)$ (or $\tilde{F}_N(f)$) is typically accompanied by geometric growth of the norm $\|a\|$ of the exact (infinite-precision) coefficient vector. Hence, whenever N is sufficiently large, one expects there to be a discrepancy between the exact coefficient vector and its numerically computed counterpart, meaning that the numerical extensions $G_N(f)$ and $\tilde{G}_N(f)$ may not exhibit the same convergence behavior. In the first half of this paper, besides showing stability, we also give a complete analysis and description of the convergence of $G_N(f)$ and $\tilde{G}_N(f)$, and discuss how this differs from that of $F_N(f)$ and $\tilde{F}_N(f)$.

We now summarize the main conclusions of the first half of the paper. Concerning stability, we have:

1. The condition numbers of the matrices A and $\tilde{A}$ of the continuous and discrete FEs are exponentially large in N (see Sect. 3.1).
2. The condition number $\kappa(F_N)$ of the exact continuous FE mapping is exponentially large in N. The condition number of the exact discrete FE mapping satisfies $\kappa(\tilde{F}_N) = 1$ for all N (see Sect. 3.4).
3. The condition number of the numerical continuous and discrete FE mappings G_N and $\tilde{G}_N$ satisfy
$$\kappa(G_N) \lesssim 1/\sqrt{\epsilon}, \qquad \kappa(\tilde{G}_N) \lesssim 1, \quad \forall N \in \mathbb{N},$$
where $\epsilon = \epsilon_{\mathrm{mach}}$ is the machine precision used (see Sect. 4.3).

To state our main conclusions regarding convergence, we first require some notation. Let $\mathcal{D}(\rho)$, $\rho \geq 1$, be a particular one-parameter family of regions in the complex plane related to Bernstein ellipses (see (2.15) and Definition 2.10), and define the Fourier extension *constant* [1, 22] by

$$E(T) = \cot^2\left(\frac{\pi}{4T}\right). \tag{1.3}$$

We now have the following:

1. Suppose that f is analytic in $\mathcal{D}(\rho^*)$ and continuous on its boundary. Then the exact continuous and discrete FEs satisfy
$$\left\|f - F_N(f)\right\|, \qquad \|f - \tilde{F}_N f\| \leq c_f \rho^{-N},$$
where $\rho = \min\{\rho^*, E(T)\}$ and c_f is proportional to $\max_{x \in \mathcal{D}(\rho)} |f(x)|$ (see Sect. 2.3).
2. For f as in 4. the errors of the numerical continuous and discrete FEs satisfy (see Sect. 4.2):
   (i) For $N \leq N_0$ (continuous) or $N \leq N_1 := 2N_0$ (discrete), where N_0 is a function-independent breakpoint depending on ϵ and T only, both $\|f - G_N(f)\|$ and $\|f - \tilde{G}_N f\|$ decay like ρ^{-N}, where ρ is as in 4.
   (ii) When $N = N_0$ or $N = N_1$, the errors
   $$\left\|f - G_{N_0}(f)\right\| \approx c_f(\sqrt{\epsilon})^{d_f}, \qquad \left\|f - \tilde{G}_{N_1}(f)\right\| \approx c_f \epsilon^{d_f},$$
   where c_f is as in 4. and $d_f = \frac{\log \rho}{\log E(T)} \in (0, 1]$.
   (iii) When $N > N_0$ or $N > N_1$, the errors decay at least superalgebraically fast down to *maximal achievable accuracies* of order $\sqrt{\epsilon}$ and ϵ, respectively. In

other words,

$$\limsup_{N\to\infty}\|f-G_N(f)\|\lesssim\sqrt{\epsilon},\qquad \limsup_{N\to\infty}\|f-\tilde{G}_N(f)\|\lesssim\epsilon.$$

Remark 1.1 In this paper we refer to several different types of convergence of an approximation $f_N \approx f$. We say that f_N converges *algebraically* fast to f at rate k if $\|f - f_N\| = \mathcal{O}(N^{-k})$ as $N \to \infty$. If $\|f - f_N\|$ decays faster than any algebraic power of N^{-1} then f_N is said to converge *superalgebraically* fast. We say that f_N converges *geometrically* fast to f if there exists a $\rho > 1$ such that $\|f - f_N\| = \mathcal{O}(\rho^{-N})$. We shall also occasionally use the term *root-exponential* to describe convergence of the form $\|f - f_N\| = \mathcal{O}(\rho^{-\sqrt{N}})$.

As we explain in Sect. 4, the reason for the disparity between the exact and numerical FEs can be traced to the fact that the system of functions $\{e^{i\frac{n\pi}{T}\cdot}\}_{n\in\mathbb{Z}}$ forms a *frame* for $L^2(-1,1)$. The inherent redundancy of this frame, i.e. the fact that any function f has infinitely many expansions in this system, leads to both the ill-conditioning in the coefficients and the differing convergence between the exact and numerical approximations F_N, $\tilde{F}_N$, and G_N, $\tilde{G}_N$, respectively.

This aside, observe that conclusion 5. asserts that the numerical continuous FE $G_N(f)$ converges geometrically fast in the regime $N < N_0$ down to an error of order $(\sqrt{\epsilon})^{d_f}$, and then at least superalgebraically fast for $N > N_0$ down to a best achievable accuracy of order $\sqrt{\epsilon}$. Note that $d_f = 1$ whenever f is analytic in $\mathcal{D}(\rho)$ with $\rho \geq E(T)$. Thus G_N approximates all sufficiently analytic functions possessing moderately small constants c_f with geometric convergence down to order $\sqrt{\epsilon}$, and this is achieved at $N = N_0$. For functions only analytic in regions $\mathcal{D}(\rho)$ with $\rho < E(T)$, or possessing large constants c_f, this accuracy is obtained after a further regime of at least superalgebraic convergence. Note that c_f is large typically when f is oscillatory or possessing boundary layers. Hence for such functions, even though they may well be entire, one usually still sees the second phase of superalgebraic convergence.

The limitation of $\sqrt{\epsilon}$ accuracy for the numerical continuous FE is undesirable. Since $\epsilon = \epsilon_{\text{mach}} \approx 10^{-16}$ in practice, this means that one cannot expect to obtain more than 7 or 8 digits of accuracy in general. The condition number is also large—specifically, $\kappa(G_N) \approx 10^8$ (see 3.)—and hence the continuous FE has limited practical value. This is in addition to $G_N(f)$ being difficult to compute in practice, since it requires calculation of $2N + 1$ Fourier integrals of f (see Sect. 2.2.1).

On the other hand, conclusion 3. shows that the discrete FE is completely stable when implemented numerically. Moreover, it possesses the same qualitative convergence behavior as the continuous FE, but with two key differences. First, the region of guaranteed geometric convergence is precisely twice as large, $N_1 = 2N_0$. Second, the maximal achievable accuracy is on the order of machine precision, as opposed to its square root (see 5.). Thus, an important conclusion of the first half of this paper is the following: it is possible to compute a numerically stable FE of any analytic function which converges at least superalgebraically fast in N (in particular, geometrically fast for all small N), and which attains close to machine accuracy for N sufficiently large.

Remark 1.2 This paper is about the discrepancy between theoretical properties of solutions to (1.1) and (1.2) and their numerical solutions when computed with standard solvers. Throughout we shall consistently use *Mathematica*'s `LeastSquares` routine in our computations, though we would like to stress that *Matlab*'s command \ gives similar results. Occasionally, to compare theoretical and numerical properties, we shall carry out computations in additional precision to eliminate the effect of round-off error. When done, this will be stated explicitly. Otherwise, it is to be assumed that all computations are carried out as described in standard precision.

1.3 Fourier Extensions from Equispaced Data

In many applications, one is faced with the problem of recovering an analytic function f to high accuracy from its values on an equispaced grid $\{f(\frac{n}{M}) : n = -M, \dots, M\}$. This problem turns out to be quite challenging. For example, the famous Runge phenomenon states that the polynomial interpolant of this data will diverge geometrically fast as $M \to \infty$ unless f is analytic in a sufficiently large region.

Numerous approaches have been proposed to address this problem, and thereby 'overcome' the Runge phenomenon (see [8, 28] for a comprehensive list). Whilst many are quite effective in practice, ill-conditioning is often an issue. This was recently explained by Platte, Trefethen and Kuijlaars in [28] (see also Sect. 5.4), wherein it was shown that any exponentially convergent method for recovering analytic functions f from equispaced data must also be exponentially ill-conditioned. As was also proved, the best possible that can be achieved by a stable method is root-exponential convergence. This profound result, most likely the first of its kind for this type of problem, places an important theoretical benchmark against which all such methods must be measured.

As we show in the first half of this paper, the numerical discrete FE is well-conditioned and has good convergence properties. Yet it relies on particular interpolation points (2.11) which are not equispaced. In the second half of this paper we consider Fourier extensions based on equispaced data. In particular, if $x_n = \frac{n}{M}$ we study the so-called *equispaced* Fourier extension

$$F_{N,M}(f) := \underset{\phi \in \mathcal{G}_N}{\operatorname{argmin}} \sum_{|n| \le M} \left| f(x_n) - \phi(x_n) \right|^2, \tag{1.4}$$

and its finite-precision counterpart $G_{N,M}(f)$.

Our primary interest shall lie with the case where $M = \gamma N$ for some $\gamma \ge 1$, i.e. where the number of points M scales linearly with N. In this case we refer to γ as the *oversampling* parameter. Observe that (1.4) results in an $(2M + 1) \times (2N + 1)$ least squares problem for the coefficients of $F_{N,M}(f)$. We shall denote the corresponding matrix by $\bar{A}$.

Our main conclusions concerning the exact equispaced FE $F_{N,M}(f)$ are as follows (see Sect. 5.2):

6. The condition number of $\bar{A}$ is exponentially large as $N, M \to \infty$ with $M \ge N$.
7. The condition number of exact equispaced FE mapping $\kappa(F_{N,\gamma N})$ is exponentially large in N whenever $M = \gamma N$ for $\gamma \ge 1$ fixed. Moreover, the approximation $F_{N,\gamma N}(f)$ suffers from a Runge phenomenon for any fixed $\gamma \ge 1$. In particular, the error $\|f - F_{N,\gamma N}(f)\|$ may diverge geometrically fast in N for certain analytic functions f.

8. The scaling $M = \mathcal{O}(N^2)$ is required to overcome the ill-conditioning and the Runge phenomenon in $F_{N,M}$. In this case, $F_{N,M}(f)$ converges at the same rate as the exact continuous FE $F_N(f)$, i.e. geometrically fast in N. Although the condition number of $\bar{A}$ remains exponentially large, the condition number of the mapping $\kappa(F_{N,M})$ is $\mathcal{O}(1)$ for this scaling.

These results lead to the following conclusion. The exact (infinite-precision) equispaced FE $F_{N,M}$ with $M = \mathcal{O}(N^2)$ attains the stability barrier of Platte, Trefethen and Kuijlaars: namely, it is well-conditioned and converges root-exponentially fast in the parameter M.

However, since the matrix $\bar{A}$ is always ill-conditioned, one expects there to be differences between the exact equispaced extension $F_{N,M}(f)$ and its numerical counterpart $G_{N,M}(f)$. In practice, one sees both differing stability and convergence behavior of $G_{N,M}(f)$, much like in the case of continuous and discrete FEs. Specifically, in Sect. 5.3 we show the following:

9. The condition number $\kappa(G_{N,\gamma N})$ satisfies

$$\kappa(G_{N,\gamma N}) \lesssim \epsilon^{-a(\gamma;T)}, \quad \forall N \in \mathbb{N},$$

where $\epsilon = \epsilon_{\text{mach}}$ is the machine precision used, and $0 < a(\gamma; T) \leq 1$ is independent of N and satisfies $a(\gamma; T) \to 0$ as $\gamma \to \infty$ for fixed T (see (5.23) for the definition of $a(\gamma; T)$).

10. The error $\|f - G_{N,\gamma N}(f)\|$ behaves as follows:
    (i) If $N < N_2$, where N_2 is a function-independent breakpoint, $\|f - G_{N,\gamma N}(f)\|$ converges or diverges exponentially fast at the same rate as $\|f - F_{N,\gamma N}(f)\|$.
    (ii) If $N_2 \leq N < N_1$, where N_1 is as introduced previously in Sect. 1.2, then $\|f - G_{N,\gamma N}(f)\|$ converges geometrically fast at the same rate as $\|f - F_N(f)\|$, where $F_N(f)$ is the exact continuous FE.
    (iii) When $N = N_1$ the error

$$\left\|f - G_{N_1,\gamma N_1}(f)\right\| \approx c_f \epsilon^{d_f - a(\gamma;T)},$$

    where c_f and d_f are as in 5. of Sect. 1.2.
    (iii) If $N > N_1$ then $\|f - G_{N,\gamma N}(f)\|$ decays at least superalgebraically fast in N down to a maximal achievable accuracy of order $\epsilon^{1-a(\gamma;T)}$.

These results show that the condition number of the numerical equispaced FE is bounded whenever $M = \gamma N$, unlike for its exact analogue. Moreover, after a (function-independent) regime of possible divergence, we witness geometric convergence of $G_{N,\gamma N}(f)$ down to a certain accuracy. As in the case of the continuous or discrete FEs, if the function f is sufficiently analytic with small constant c_f then the convergence effectively stops at this point. If not, we witness a further regime of guaranteed superalgebraic convergence. But in both cases, the maximal achievable accuracy is of order $\epsilon^{1-a(\gamma;T)}$, which, since $a(\gamma; T) \to 0$ as $\gamma \to \infty$, can be made arbitrarily close to ϵ by increasing γ. Note that doing this both improves the condition number of the numerical equispaced FE and yields a less severe rate of exponential divergence in the region $N < N_2$. As we show via numerical computation of the relevant constants, double oversampling $\gamma = 2$ with $T = 2$ gives perfectly adequate results in most cases.

The main conclusion of this analysis is that numerical equispaced FEs, unlike their exact counterparts, are able to circumvent the stability barrier of Platte, Trefethen and Kuijlaars to an extent (see Sect. 5.4 for a more detailed discussion). Specifically, the numerical FE $F_{N,\gamma N}$ has a bounded condition number, and for all sufficiently analytic functions—namely, those analytic in the region $\mathcal{D}(E(T))$—the convergence is geometric down to a finite accuracy of order $c_f \epsilon^{1-a(\gamma;T)}$. This latter observation, namely the fact that the maximal accuracy is nonzero, is precisely the reason why the stability theorem, which requires geometric convergence for all N, does not apply. On the other hand, for all other analytic functions (or those possessing large constants c_f) the convergence is at least superalgebraic for $N > N_1$ down to roughly $\epsilon^{1-a(\gamma;T)}$; again not in contradiction with the theorem. Importantly, one never sees divergence of the numerical FE after the finite breakpoint N_2.

For this reason, we conclude that equispaced FEs are an attractive method for approximations from equispaced data. To further support this conclusion we also remark that although the primary concern of this paper is analytic functions, equispaced FEs are also applicable to functions of finite regularity. In this case, one witnesses algebraic convergence, with the precise order depending solely on the degree of smoothness (see Theorem 2.9).

1.4 Relation to Previous Work

One-dimensional FEs for overcoming the Gibbs and Runge phenomena were studied in [5] and [8], and applications to surface parametrizations considered in [12]. Analysis of the convergence of the exact continuous and discrete FEs was presented by Huybrechs in [22] and Adcock and Huybrechs in [1]. The issue of resolution power was also addressed in the latter. The content of the first half of this paper, namely analysis of exact/numerical FEs, follows on directly from this work.

A different approach to FEs, known as the FC–Gram method, was introduced in [26]. This approach forms a central part of an extremely effective method for solving PDEs in complex geometries [2, 11]. For previous work on using FEs for PDE problems (so-called Fourier *embeddings*) see [6, 27].

Equispaced FEs of the form studied in this paper were first independently considered by Boyd [5] and Bruno [10], and later by Bruno et al. [12]. In particular, Boyd [5] describes the use of truncated singular value decompositions (SVDs) to compute equispaced FEs, and gives extensive numerical experiments (see also [8]). Bruno focuses on the use of Fourier extensions (also called *Fourier continuations* in the above references) for the description of complicated smooth surfaces. He suggested in [10] a weighted least squares to obtain a smooth extension for this purpose, with numerical evidence supporting convergence results in [12]. Most recently Lyon has presented an analysis of equispaced FEs computed using truncated SVDs [24]. In particular, numerical stability and convergence (down to close to machine precision) were shown. In Sect. 5.3 we discuss this work in more detail (see, in particular, Remark 5.10), and give further insight into some of the questions raised in [24].

1.5 Outline of the Paper

The outline of the remainder of this paper is as follows. In Sect. 2 we recap properties of the continuous and discrete FEs from [1, 22], including convergence and how to choose the extension parameter T. Ill-conditioning of the coefficient map is proved in Sect. 3, and in Sect. 4 we consider the stability of the numerical extensions and their convergence. Finally, in Sect. 5 we consider the case of equispaced FEs.

A comprehensive list of symbols is given at the end of the paper.

2 Fourier Extensions

In this section we introduce FEs, and recap salient important aspects of [1, 22].

2.1 Two Interpretations of Fourier Extensions

There are two important interpretations of FEs which inform their approximation properties and their stability, respectively. These are described in the next two sections.

2.1.1 Fourier Extensions as Polynomial Approximations

The space $\mathcal{G}_N$ can be decomposed as $\mathcal{G}_N = \mathcal{C}_N \oplus \mathcal{S}_N$, where

$$\mathcal{C}_N = \operatorname{span}\left\{\cos\frac{n\pi}{T}x : n = 0, \dots, N\right\}, \qquad \mathcal{S}_N = \operatorname{span}\left\{\sin\frac{n\pi}{T}x : n = 1, \dots, N\right\},$$

consist of even and odd functions, respectively. Likewise, for f we have

$$f(x) = f_{\mathrm{e}}(x) + f_{\mathrm{o}}(x), \qquad f_{\mathrm{e}}(x) = \frac{1}{2}\big[f(x) + f(-x)\big],\ f_{\mathrm{o}}(x) = \frac{1}{2}\big[f(x) - f(-x)\big],$$

and for any FE f_N of f:

$$f_N = f_{\mathrm{e},N} + f_{\mathrm{o},N}, \quad f_{\mathrm{e},N} \in \mathcal{C}_N,\ f_{\mathrm{o},N} \in \mathcal{S}_N. \tag{2.1}$$

Throughout this paper we shall use the notation f_N to denote an arbitrary FE of f when not wishing to specify its particular construction. From (2.1), it follows that the problem of approximating f via a FE f_N decouples into two problems $f_{\mathrm{e},N} \approx f_{\mathrm{e}}$ and $f_{\mathrm{o},N} \approx f_{\mathrm{o}}$ in the subspaces $\mathcal{C}_N$ and $\mathcal{S}_N$, respectively, on the half-interval [0, 1].

Let us define the mapping $y = y(x) : [0, 1] \to [c(T), 1]$ by $y = \cos\frac{\pi}{T}x$, where $c(T) = \cos\frac{\pi}{T}$. The functions $\cos\frac{n\pi}{T}x$ and $\sin\frac{(n+1)\pi}{T}x / \sin\frac{\pi}{T}x$ are algebraic polynomials of degree n in y. Therefore $\mathcal{C}_N$ and $\mathcal{S}_N$ are (up to multiplication by $\sin\frac{\pi}{T}x$ for the latter) the subspaces $\mathbb{P}_N$ and $\mathbb{P}_{N-1}$ of polynomials of degree N and $N-1$, respectively, in the transformed variable y. Letting

$$g_1(y) = f_{\mathrm{e}}(x), \qquad g_2(y) = \frac{f_{\mathrm{o}}(x)}{\sin\frac{\pi}{T}x},$$

$$g_{1,N}(y) = f_{\mathrm{e},N}(x), \qquad g_{2,N}(y) = \frac{f_{\mathrm{o},N}(x)}{\sin \frac{\pi}{T} x},$$

with $g_{1,N}(y) \in \mathbb{P}_N$ and $g_{2,N}(y) \in \mathbb{P}_{N-1}$, we conclude that the FE approximation f_N in the variable x is completely equivalent to two polynomial approximations in the transformed variable $y \in [c(T), 1]$.

This fact is central to the analysis of FEs. It allows one to use the rich literature on polynomial approximations to determine the theoretical behavior of the continuous and discrete FEs (see Sect. 2.3).

Remark 2.1 The interpretation of f_N in terms of polynomials is solely for the purposes of analysis. We always perform computations in the x-domain using the standard trigonometric basis for $\mathcal{G}_N$ (see Sect. 2.2).

The interval $[c(T), 1] \subseteq (-1, 1]$ is not standard. It is thus convenient to map it affinely to $[-1, 1]$. Let

$$z := z(y) = 2\frac{y - c(T)}{1 - c(T)} - 1 \in [-1, 1].$$

Observe that $y = y(z) = c(T) + \frac{1-c(T)}{2}(z + 1)$. Let $m : [0, 1] \to [-1, 1]$ be the mapping $x \mapsto z$, i.e.

$$z = m(x) = 2\frac{\cos \frac{\pi}{T} x - c(T)}{1 - c(T)} - 1. \tag{2.2}$$

Note that $x = m^{-1}(z) = \frac{T}{\pi} \arccos[c(T) + \frac{1-c(T)}{2}(z + 1)]$. If we now define

$$h_i(z) = g_i\bigl(y(z)\bigr), \quad i = 1, 2, \tag{2.3}$$

then the FE f_N is equivalent to the two polynomial approximations

$$\begin{aligned} h_{1,N}(z) &= g_{1,N}\bigl(y(z)\bigr) = f_{\mathrm{e},N}\bigl(m^{-1}(z)\bigr), \\ h_{2,N}(z) &= g_{2,N}\bigl(y(z)\bigr) = \frac{f_{\mathrm{o},N}(m^{-1}(z))}{\sin(\frac{\pi}{T} m^{-1}(z))}, \end{aligned} \tag{2.4}$$

of degree N and $N - 1$ respectively in the new variable $z \in [-1, 1]$.

2.1.2 Fourier Extensions as Frame Approximations

Definition 2.2 Let H be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. A set $\{\phi_n\}_{n=1}^{\infty} \subseteq \mathrm{H}$ is a frame for H if (i) $\mathrm{span}\{\phi_n\}_{n=1}^{\infty}$ is dense in H and (ii) there exist $c_1, c_2 > 0$ such that

$$c_1 \|f\|^2 \le \sum_{n=1}^{\infty} \bigl|\langle f, \phi_n \rangle\bigr|^2 \le c_2 \|f\|^2, \quad \forall f \in \mathrm{H}. \tag{2.5}$$

If $c_1 = c_2$ then $\{\phi_n\}_{n=1}^{\infty}$ is referred to as a tight frame.

Introduced by Duffin and Schaeffer [16], frames are vitally important in signal processing [14]. Note that all orthonormal, indeed Riesz, bases are frames, but a frame need not be a basis. In fact, frames are typically *redundant*: any element $f \in \mathrm{H}$ may well have infinitely many representations of the form $f = \sum_{n=1}^{\infty} \alpha_n \phi_n$ with coefficients $\{\alpha_n\}_{n=1}^{\infty} \in l^2(\mathbb{N})$.

The relevance of frames to Fourier extensions is due to the following observation:

Lemma 2.3 [1] *The set* $\{\frac{1}{\sqrt{2T}} \mathrm{e}^{\mathrm{i}\frac{n\pi}{T}x}\}_{n\in\mathbb{Z}}$ *is a tight frame for* $\mathrm{L}^2(-1,1)$ *with* $c_1 = c_2 = 1$.

Note that $\{\frac{1}{\sqrt{2T}} \mathrm{e}^{\mathrm{i}\frac{n\pi}{T}x}\}_{n\in\mathbb{Z}}$ is an orthonormal basis for $\mathrm{L}^2(-T,T)$: it is precisely the standard Fourier basis on $[-T,T]$. However, it forms only a frame when considered as a subset of $\mathrm{L}^2(-1,1)$. This fact means that ill-conditioning may well be an issue in numerical algorithms for computing FEs, due to the possibility of redundancies. As it happens, it is trivial to see that the set $\{\frac{1}{\sqrt{2T}} \mathrm{e}^{\mathrm{i}\frac{n\pi}{T}x}\}_{n\in\mathbb{Z}}$ is redundant:

Lemma 2.4 *Let* $f \in \mathrm{L}^2(-1,1)$ *be arbitrary, and suppose that* $\tilde{f} \in \mathrm{L}^2(-T,T)$ *is such that* $f = \tilde{f}$ *a.e. on* $[-1,1]$. *If* $\phi_n(x) = \frac{1}{\sqrt{2T}} \mathrm{e}^{\mathrm{i}\frac{n\pi}{T}x}$ *and* $\alpha_n = \langle \tilde{f}, \phi_n \rangle_{[-T,T]}$, *then*

$$f = \sum_{n\in\mathbb{Z}} \alpha_n \phi_n \quad a.e. \tag{2.6}$$

In particular, there are infinitely many sequences $\{\alpha_n\}_{n\in\mathbb{Z}} \in l^2(\mathbb{Z})$ *for which* $f = \sum_{n\in\mathbb{Z}} \alpha_n \phi_n$.

Proof The sum $\sum_{n\in\mathbb{Z}} \alpha_n \phi_n$ is the Fourier series of $\tilde{f}$ on $[-T,T]$. Thus it coincides with $\tilde{f}$ a.e. on $[-T,T]$, and hence f when restricted to $[-1,1]$. Since there are infinitely many possible $\tilde{f}$, each giving rise to a different sequence $\{\alpha_n\}_{n\in\mathbb{Z}}$, the result now follows. □

This lemma is valid for arbitrary $f \in \mathrm{L}^2(-1,1)$. When f has higher regularity—say $f \in \mathrm{H}^k(-1,1)$, where $\mathrm{H}^k(-1,1)$ is the kth standard Sobolev space on $(-1,1)$—it is useful to note that there exist extensions $\tilde{f}$ with the same regularity on the torus $\mathbb{T} = [-T,T)$. This is the content of the next result. For convenience, given a domain I, we now write $\|\cdot\|_{\mathrm{H}^k(I)}$ for the standard norm on $\mathrm{H}^k(I)$:

Lemma 2.5 *Let* $f \in \mathrm{H}^k(-1,1)$ *for some* $k \in \mathbb{N}$. *Then there exists an extension* $\tilde{f} \in \mathrm{H}^k(\mathbb{T})$ *of* f *satisfying* $\|\tilde{f}\|_{\mathrm{H}^k(\mathbb{T})} \leq c_k(T)\|f\|_{\mathrm{H}^k(-1,1)}$, *where* $c_k(T) > 0$ *is independent of* f. *Moreover,* $f = \sum_{n\in\mathbb{Z}} \alpha_n \phi_n$, *where* $\alpha_n = \langle \tilde{f}, \phi_n \rangle_{[-T,T]}$ *satisfies* $\alpha_n = \mathcal{O}(n^{-k})$ *as* $|n| \to \infty$.

Proof The first part of the lemma follows directly from the proof of Theorem 2.1 in [1]. The second follows from integrating by parts k times and using the fact that $\tilde{f}$ is periodic. □

This lemma, which shall be important later when studying numerical FEs, states that there exist representations of f in the frame $\{\frac{1}{\sqrt{2T}}\mathrm{e}^{\mathrm{i}\frac{n\pi}{T}x}\}_{n\in\mathbb{Z}}$ that have nice (i.e. rapidly decaying) coefficients and which cannot grow large on the extended region $[-T,T]$.

2.2 The Continuous and Discrete Fourier Extensions

We now describe the two types of FEs we consider in the first part of this paper.

2.2.1 The Continuous Fourier Extension

The continuous FE of $f\in\mathrm{L}^2(-1,1)$, defined by (1.1), is the orthogonal projection onto $\mathcal{G}_N$. Computation of this extension involves solving a linear system. Let us write $F_N(f)=\sum_{n=-N}^{N}a_n\phi_n$ with unknowns $\{a_n\}_{n=-N}^{N}$. If $a=(a_{-N},\dots,a_N)^\top$ and $b=(b_{-N},\dots,b_N)^\top$, where

$$b_n=\langle f,\phi_n\rangle=\int_{-1}^{1}f(x)\overline{\phi_n(x)}\,\mathrm{d}x,\quad n=-N,\dots,N, \tag{2.7}$$

and $A\in\mathbb{C}^{(2N+1)\times(2N+1)}$ is the matrix with (n,m)th entry

$$A_{n,m}=\langle\phi_m,\phi_n\rangle=\int_{-1}^{1}\phi_m(x)\overline{\phi_n(x)}\,\mathrm{d}x,\quad n,m=-N,\dots,N, \tag{2.8}$$

then a is the solution of the linear system $Aa=b$. We refer to the values $\{a_n\}_{n=-N}^{N}$ as the *coefficients* of the FE $F_N(f)$. Note that the matrix A is a Hermitian positive-definite, Toeplitz matrix with $A_{n,m}=A_{n-m}$, where $A_0=\frac{1}{T}$ and $A_n=\frac{\sin\frac{n\pi}{T}}{n\pi}$ otherwise. In fact, A coincides with the so-called *prolate* matrix [31, 33]. We shall discuss this connection further in Sect. 3.2.

For later use, we also note the following characterization of $F_N(f)$:

Proposition 2.6 [1, 22] *Let $F_N(f)$ be the continuous FE* (1.1) *of a function f, and let $h_i(z)$ and $h_{i,N}(z)$ be given by* (2.3) *and* (2.4), *respectively* (*i.e. the symmetric and anti-symmetric parts of f and f_N with the coordinate transformed from the trigonometric argument x to the polynomial argument z*). *Then $h_{1,N}(z)$ and $h_{2,N}(z)$ are the truncated expansions of $h_1(z)$ and $h_2(z)$, respectively, in polynomials orthogonal with respect to the weight functions*

$$w_1(z)=\big[(1-z)\big(z-m(T)\big)\big]^{-\frac{1}{2}},\qquad w_2(z)=\big[(1-z)\big(z-m(T)\big)\big]^{\frac{1}{2}},\quad z\in[-1,1], \tag{2.9}$$

where $m(T)=1-2\operatorname{cosec}^2(\frac{\pi}{2T})<-1$. In other words, $h_{i,N}(z)$, $i=1,2$, is the orthogonal projection of $h_i(z)$ onto $\mathbb{P}_{N+1-i}$ with respect to the weighted inner product $\langle\cdot,\cdot\rangle_{w_i}$ with weight function w_i.

2.2.2 The Discrete Fourier Extension

The discrete FE $\tilde{F}_N(f)$ is defined by (1.2). To use this extension it is first necessary to choose nodes $\{x_n\}_{n=-N}^{N}$. This question was considered in [1], and a solution was obtained by exploiting the characterization of FEs as polynomial approximations in the transformed variable z.

A good system of nodes for polynomial interpolation is given by the Chebyshev nodes

$$z_n = \cos\left(\frac{(2n+1)\pi}{2N+2}\right), \quad n = 0, \ldots, N. \tag{2.10}$$

Mapping these back to the x-variable and symmetrizing about $x = 0$ leads to the so-called *mapped symmetric Chebyshev* nodes

$$x_n = -x_{-n-1} = \frac{T}{\pi} \arccos\left[\frac{1}{2}\big(1 - c(T)\big)\cos\left(\frac{(2n+1)\pi}{2N+2}\right) + \frac{1}{2}\big(1 + c(T)\big)\right], \quad n = 0, \ldots, N. \tag{2.11}$$

This gives a set of $2N+2$ nodes. Therefore, rather than (1.2), we define the discrete FE by

$$\tilde{F}_N(f) := \underset{\phi \in \mathcal{G}'_N}{\operatorname{argmin}} \sum_{n=-N-1}^{N} \left|f(x_n) - \phi(x_n)\right|^2, \tag{2.12}$$

from now on, where $\mathcal{G}'_N = \mathcal{C}_N \oplus \mathcal{S}_{N+1}$. Exploiting the relation between FEs and polynomial approximations once more, we now obtain the following:

Proposition 2.7 *Let $f_N = \tilde{F}_N(f) \in \mathcal{G}'_N$ be the discrete FE* (2.12) *based on the nodes* (2.11), *and let $h_i(z)$ and $h_{i,N}(z) \in \mathbb{P}_N$ be given by* (2.3) *and* (2.4), *respectively. Then $h_{i,N}(z)$, $i = 1, 2$ is the Nth degree polynomial interpolant of $h_i(z)$ at the Chebyshev nodes* (2.10).

Write $\phi_n(x) = \cos\frac{n\pi}{T}x$, $\phi_{-(n+1)}(x) = \sin\frac{n+1}{T}\pi x$, $n \in \mathbb{N}$, and let $\tilde{F}_N(f)(x) = \sum_{n=-N-1}^{N} a_n \phi_n(x)$. If $a = (a_{-N-1}, \ldots, a_N)^{-\mathrm{T}}$ and $\tilde{A} \in \mathbb{R}^{(2N+2)\times(2N+2)}$ has (n,m)th entry

$$\tilde{A}_{n,m} = \sqrt{\frac{\pi}{N+1}}\phi_m(x_n), \quad n, m = -N-1, \ldots, N, \tag{2.13}$$

then we have $\tilde{A}a = \tilde{b}$, where $\tilde{b} = (\tilde{b}_{-N-1}, \ldots, \tilde{b}_N)^\top$ and $\tilde{b}_n = \sqrt{\frac{\pi}{N+1}} f(x_n)$.

The following lemma concerning the matrix $\tilde{A}$ will prove useful in what follows:

Lemma 2.8 [1] *The matrix $A_W = (\tilde{A})^* \tilde{A}$ has entries*

$$\langle \phi_n, \phi_m \rangle_W := \int_{-1}^{1} \phi_n(x)\phi_m(x)W(x)\,\mathrm{d}x, \quad n, m = -N-1, \ldots, N,$$

where W is the positive, integrable weight function given by $W(x) = \frac{\sqrt{2}\pi}{T}\frac{\cos\frac{\pi}{2T}x}{\sqrt{\cos\frac{\pi}{T}x-\cos\frac{\pi}{T}}}$.

This lemma implies that the left-hand side of the normal equations of the discrete FE are the equations of a continuous FE based on the weighted least-squares minimization with weight function W.

2.3 Convergence of Exact Fourier Extensions

A detailed analysis of the convergence of the exact continuous FE, which we now recap, was carried out in [1, 22]. We commence with the following theorem:

Theorem 2.9 [1] *Suppose that* $f \in \mathrm{H}^k(-1,1)$ *for some* $k \in \mathbb{N}$ *and that* $T > 1$. *If* $F_N(f)$ *is the continuous FE of* f *defined by* (1.1), *then*

$$\|f - F_N(f)\| \le c_k(T)N^{-k}\|f\|_{\mathrm{H}^k(-1,1)}, \quad \forall N \in \mathbb{N}, \tag{2.14}$$

where $c_k(T) > 0$ *is independent of* f *and* N.

This theorem confirms *algebraic* convergence of $F_N(f)$ whenever the approximated function f has finite degrees of smoothness, and *superalgebraic* convergence, i.e. faster than any fixed algebraic power of N^{-1}, whenever $f \in \mathrm{C}^\infty[-1,1]$.

Suppose now that f is analytic. Although superalgebraic convergence is guaranteed by Theorem 2.9, it transpires that the convergence is actually geometric. This is a direct consequence of the interpretation of the $F_N(f)$ as the sum of two polynomial expansions in the transformed variable z (Proposition 2.6). To state the corresponding theorem, we first require the following definition:

Definition 2.10 The Bernstein ellipse $\mathcal{B}(\rho) \subseteq \mathbb{C}$ of index $\rho \ge 1$ is given by

$$\mathcal{B}(\rho) = \left\{\frac{1}{2}\left(\rho^{-1}\mathrm{e}^{\mathrm{i}\theta} + \rho\mathrm{e}^{-\mathrm{i}\theta}\right) : \theta \in [-\pi,\pi]\right\}.$$

Given a compact region bounded by the Bernstein ellipse $\mathcal{B}(\rho)$, we shall write

$$\mathcal{D}(\rho) \subseteq \mathbb{C} \tag{2.15}$$

for its image in the complex x-plane under the mapping $x = m^{-1}(z)$, where m is as in (2.2).

Theorem 2.11 [1, 22] *Suppose that* f *is analytic in* $\mathcal{D}(\rho^*)$ *and continuous on its boundary. Then* $\|f - F_N(f)\|_\infty \le c_f \rho^{-N}$, *where* $\rho = \min\{\rho^*, E(T)\}$, $c_f > 0$ *is proportional to* $\max_{x\in\mathcal{D}(\rho)}|f(x)|$, *and* $E(T)$ *is as in* (1.3).

Proof A full proof was given in [1, Theorem 2.3]. The expansion g_N of an analytic function g in a system of orthogonal polynomials with respect to some integrable weight function satisfies $\|g - g_N\|_\infty \le c_g\rho^{-N}$, where c_g is proportional to

$\max_{z\in\mathcal{B}(\rho)}|g(z)|$ [30]. In view of Proposition 2.6, it remains only to determine the maximal parameter ρ of Bernstein ellipse $\mathcal{B}(\rho)$ within which $h_1(z)$ and $h_2(z)$ are analytic.

The mapping $z=m(x)$ introduces a square-root type singularity into the functions $h_i(z)$ at the point $z=m(T)<-1$. Hence the maximal possible value of the parameter ρ satisfies

$$\frac{1}{2}\left(\rho+\rho^{-1}\right)=-m(T). \tag{2.16}$$

Observe that if $\psi(t)=t+\sqrt{t^2-1}$ then

$$\psi\bigl(m(T)\bigr)=E(T). \tag{2.17}$$

Thus, since $\rho>1$, the solution to (2.16) is precisely $\rho=E(T)$. Conversely, any singularity of f introduces a singularity of $h_i(z)$, which also limits this value. Hence we obtain the stated minimum. □

Theorem 2.11 shows that if f is analytic in a sufficiently large region (for example, if f is entire) then the rate of geometric convergence is precisely $E(T)$. Recall that the parameter T can be chosen by the user. In the next section we consider the effect of different choices of T.

Remark 2.12 Although Theorems 2.9 and 2.11 are stated for $F_N(f)$, they also hold for the discrete FE $\tilde{F}_N(f)$, since the latter is equivalent to a sum of Chebyshev interpolants (Proposition 2.7).

2.4 The Choice of T

Note that $E(T)\sim 1+\pi(T-1)$ as $T\to 1^+$ and $E(T)\sim\frac{16}{\pi^2}T^2$ when $T\to\infty$. Thus, small T leads to a slower rate of geometric convergence, whereas large T gives a faster rate. As discussed in [1], however, a larger value of T leads to a worse resolution power, meaning that more degrees of freedom are required to resolve oscillatory behavior. On the other hand, setting T sufficiently close to 1 yields a resolution power that is arbitrarily close to optimal.

In [1] a number of fixed values of T were used in numerical experiments. These typically give good results, with small values of T being particularly well suited to oscillatory functions. Another approach for choosing T was also discussed. This involves letting

$$T=T(N;\epsilon_{\mathrm{tol}})=\frac{\pi}{4}\left(\arctan\left((\epsilon_{\mathrm{tol}})^{\frac{1}{2N}}\right)\right)^{-1}, \tag{2.18}$$

where $\epsilon_{\mathrm{tol}}\ll 1$ is some fixed tolerance (note that this is very much related to the Kosloff Tal-Ezer map in spectral methods for PDEs [4, 23]—see [1] for a discussion). This choice of T, which now depends on N, is such that $E(T)^{-N}=\epsilon_{\mathrm{tol}}$. Although this limits the best achievable accuracy of the FE with this approach to $\mathcal{O}(\epsilon_{\mathrm{tol}})$, setting $\epsilon_{\mathrm{tol}}=10^{-14}$ is normally sufficient in practice. Numerical experiments in [1] indicate that this works well, especially for oscillatory functions. In fact, since

$$T(N;\epsilon_{\mathrm{tol}}) \sim 1 - \frac{\log(\epsilon_{\mathrm{tol}})}{\pi N} + \mathcal{O}\big(N^{-2}\big), \quad N \to \infty, \tag{2.19}$$

this approach has formally optimal resolution power.

Remark 2.13 The strategy (2.18) is particularly good for oscillatory problems. However, if this is not a concern, a practical choice appears to be $T = 2$. In this case, the FE has a particular symmetry that can be exploited to allow for its efficient computation in only $\mathcal{O}(N(\log N)^2)$ operations [25].

3 Condition Numbers of Exact Fourier Extensions

The redundancy of the frame $\{\frac{1}{\sqrt{2T}}\mathrm{e}^{\mathrm{i}\frac{n\pi}{T}\cdot}\}_{n\in\mathbb{Z}}$ means that the matrices associated with the continuous and discrete FEs are ill-conditioned. We next derive bounds for the condition number of these matrices. The spectrum of A is considered further in Sect. 3.2, and the condition numbers of the FE mappings $f \mapsto F_N(f)$ and $f \mapsto \tilde{F}_N(f)$ are discussed in Sect. 3.4.

3.1 The Condition Numbers of the Continuous and Discrete FE Matrices

Theorem 3.1 *Let A be the matrix (2.8) of the continuous FE. Then the condition number of A is $\mathcal{O}(E(T)^{2N})$ for large N. Specifically, the maximal and minimal eigenvalues satisfy*

$$T^{-1} \le \lambda_{\max}(A) \le 1, \qquad c_1(T)N^{-3}E(T)^{-2N} \le \lambda_{\min}(A) \le c_2(T)N^2E(T)^{-2N}, \tag{3.1}$$

where $c_1(T)$ and $c_2(T)$ are positive constants with $c_1(T), c_2(T) = \mathcal{O}(1)$ as $T \to 1^+$.

Proof It is a straightforward exercise to verify that

$$\begin{aligned}\lambda_{\min}(A) &= \min_{\phi\in\mathcal{G}_N}\big\{\|\phi\|^2 : \|\phi\|_{[-T,T]} = 1\big\},\\ \lambda_{\max}(A) &= \max_{\phi\in\mathcal{G}_N}\big\{\|\phi\|^2 : \|\phi\|_{[-T,T]} = 1\big\}.\end{aligned} \tag{3.2}$$

Using the fact that $\|\phi\| \le \|\phi\|_{[-T,T]}$, we first notice that $\lambda_{\max}(A) \le 1$. On the other hand, setting $\phi = \frac{1}{\sqrt{2T}}$, we find that $\lambda_{\max}(A) \ge T^{-1}$, which completes the result for $\lambda_{\max}(A)$.

We now consider $\lambda_{\min}(A)$. Recall that any $\phi \in \mathcal{G}_N$ can be decomposed into even and odd parts ϕ_{e} and ϕ_{o}, with each function corresponding to a polynomial in the transformed variable z. Hence,

$$\lambda_{\min}(A) = \min_{\substack{\phi\in\mathcal{G}_N\\ \phi\neq 0}} \left\{ \frac{\|\phi\|^2}{\|\phi\|^2_{[-T,T]}} \right\}$$

$$= \min_{\substack{p_1\in\mathbb{P}_N, p_2\in\mathbb{P}_{N-1}\\ \|p_1\|+\|p_2\|\neq 0}} \left\{ \frac{\|p_1\|^2_{w_1} + \|p_2\|^2_{w_2}}{\|p_1\|^2_{w_1,[m(T),1]} + \|p_2\|^2_{w_2,[m(T),1]}} \right\}, \tag{3.3}$$

where w_i, $i = 1, 2$, is given by (2.9). Since the weight function w_i is integrable, we have

$$\|p_i\|_{w_i,[m(T),1]} \le \sqrt{C_i(T)}\|p_i\|_{\infty,[m(T),1]}, \quad i = 1, 2, \tag{3.4}$$

where $C_i(T) = \int_{m(T)}^1 \mathrm{d}w_i$, $i = 1, 2$. Moreover, by Remez's inequality,

$$\|p\|_{\infty,[m(T),1]} \le \|T_N\|_{\infty,[m(T),1]}\|p\|_\infty, \quad \forall p \in \mathbb{P}_N,$$

where $T_N \in \mathbb{P}_N$ is the Nth Chebyshev polynomial. Since T_N is monotonic outside $[-1, 1]$, we have $\|T_N\|_{\infty,[m(T),1]} = |T_N(m(T))|$. Moreover, due to the formula

$$T_N(x) = \frac{1}{2}\left[\left(x - \sqrt{x^2-1}\right)^n + \left(x + \sqrt{x^2-1}\right)^n\right],$$

an application of (2.17) gives

$$\|T_N\|_{\infty,[m(T),1]} = \frac{1}{2}\left[E(T)^N + E(T)^{-N}\right] < E(T)^N, \quad \forall N \in \mathbb{N}, T > 1. \tag{3.5}$$

Next we note that $w_1(z) \ge D_1(T)$ and $w_2(z) \ge D_2(T)\sqrt{1-z^2}$, $\forall z \in [-1, 1]$, for positive constants $D_1(T)$ and $D_2(T)$. Moreover, there exist constants $d_1, d_2 > 0$ independent of T such that

$$\|p\|_\infty \le d_1 N\|p\|, \qquad \|p\|_\infty \le d_2 N^{\frac{3}{2}}\|p\|_v, \quad p \in \mathbb{P}_N,$$

where $v(z) = \sqrt{1-z^2}$ (this follows from expanding p in orthonormal polynomials $\{p_n\}_{n\in\mathbb{N}}$ on $[-1, 1]$ corresponding to the weight function $w(z) = 1$, i.e. Legendre polynomials, or $w(z) = v(z)$, i.e. Chebyshev polynomials of the second kind, and using the known estimate $\|p_n\|_\infty = \mathcal{O}(n^{\frac{1}{2}})$ for the former and $\|p_n\|_\infty = \mathcal{O}(n^{\frac{3}{2}})$ for the latter [3, Chapt. X]). Therefore

$$\|p\|_\infty \le \frac{d_i}{\sqrt{D_i(T)}} N^{\frac{1+i}{2}}\|p\|_{w_i}, \quad \forall p \in \mathbb{P}_N, \ i = 1, 2. \tag{3.6}$$

Substituting (3.4), (3.5) and (3.6) into (3.3) now gives

$$\lambda_{\min}(A) \ge \frac{1}{\max\{C_1(T)/D_1(T), C_2(T)/D_2(T)\}} N^{-3}E(T)^{-2N},$$

which gives the lower bound in (3.1).

For the upper bound, we set $p_2 = 0$ and $p_1 = T_N$ in (3.3) to give

$$\lambda_{\min}(A) \leq \frac{\|T_N\|_{w_1}^2}{\|T_N\|_{w_1,[m(T),1]}^2} \leq \frac{C_1(T)}{\|T_N\|_{w_1,[m(T),1]}^2}. \tag{3.7}$$

Using (3.5) we note that $\|T_N\|_{\infty,[m(T),1]} \geq \frac{1}{2}E(T)^N$. Recall also that $\|p\|_\infty \leq d_1 N\|p\|$, $\forall p \in \mathbb{P}_N$. Scaling this inequality to the interval $[m(T), 1]$ now gives

$$\|p\|_{\infty,[m(T),1]} \leq d_1\sqrt{\frac{2}{1-m(T)}}N\|p\|_{[m(T),1]} = \sqrt{C_3(T)}N\|p\|_{[m(T),1]}.$$

Note also that $w_1(z) \geq D_3(T)$, $\forall z \in [m(T), 1]$. Therefore,

$$\begin{aligned}\|T_N\|_{w_1,[m(T),1]} &\geq \sqrt{D_3(T)}\|T_N\|_{[m(T),1]}\\ &\geq \frac{\sqrt{D_3(T)}}{\sqrt{C_3(T)}N}\|T_N\|_{\infty,[m(T),1]}\\ &\geq \frac{\sqrt{D_3(T)}}{2\sqrt{C_3(T)}N}E(T)^N.\end{aligned}$$

Substituting this into (3.7) now gives the result. □

We now consider the case of the discrete FE:

Theorem 3.2 *Let $\tilde{A}$ be the matrix (2.13) of the discrete FE. Then the condition number of $\tilde{A}$ is $\mathcal{O}(E(T)^N)$ for large N. Specifically, the maximal and minimal singular values of $\tilde{A}$ satisfy*

$$\begin{aligned}&c_1(T) \leq \sigma_{\max}(\tilde{A}) \leq c_2(T)N^{\frac{3}{2}},\\ &d_1(T)N^{-\frac{3}{2}}E(T)^{-N} \leq \sigma_{\min}(\tilde{A}) \leq d_2(T)N^{\frac{5}{2}}E(T)^{-N},\end{aligned} \tag{3.8}$$

where $c_1(T), c_2(T), d_1(T), d_2(T)$ are positive constants that are $\mathcal{O}(1)$ as $T \to 1^+$.

Proof Using Lemma 2.8, the values $\sigma_{\min}^2(\tilde{A})$ and $\sigma_{\max}^2(\tilde{A})$ may be expressed as in (3.2) (with $\|\cdot\|$ replaced by $\|\cdot\|_W$). Note that $W(0)\|\phi\|^2 \leq \|\phi\|_W^2 \leq \|\phi\|_\infty^2 \int_{-1}^1 \mathrm{d}W$. It is a straightforward exercise (using the bound (3.6) and the fact that ϕ can be expressed as the sum of two polynomials) to show that $\|\phi\|_\infty \leq C_1(T)N^{\frac{3}{2}}\|\phi\|$, where $C_1(T) = \mathcal{O}(1)$ as $T \to 1^+$. Thus we obtain

$$W(0)\frac{\|\phi\|^2}{\|\phi\|_{[-T,T]}^2} \leq \frac{\|\phi\|_W^2}{\|\phi\|_{[-T,T]}^2} \leq \left(C_1(T)^2\int_{-1}^1 \mathrm{d}W\right)N^3\frac{\|\phi\|^2}{\|\phi\|_{[-T,T]}^2}.$$

The result now follows immediately from the bounds (3.1). □

Theorems 3.1 and 3.2 demonstrate that the condition numbers of the continuous and discrete FE matrices grow exponentially in N. This establishes conclusion 1. of Sect. 1.

Remark 3.3 Although exponentially large, the matrix of the discrete FE is substantially less poorly conditioned than that of the continuous FE. In particular, the condition number is of order $E(T)^N$ as opposed to $E(T)^{2N}$. This can be understood using Lemma 2.8. The normal form $A_W = (\tilde{A})^* \tilde{A}$ of the discrete FE matrix is a continuous FE matrix with respect to the weight function A_W. Hence $\kappa(\tilde{A}) = \sqrt{\kappa(A_W)} \approx \sqrt{\kappa(A)} \approx E(T)^N$. As we shall see later, this property also translates into superior performance of the numerical discrete FE over its continuous counterpart (see Sect. 4.2).

Since the constants in Theorems 3.1 and 3.2 are bounded as $T \to 1^+$, this allows one also to determine the condition number in the case that $T \to 1^+$ as $N \to \infty$ (see Sect. 2.4). In particular, if T is given by (2.18), then $\kappa(A)$ and $\kappa(\tilde{A})$ are (up to possible small algebraic factors in N) of order $(\epsilon_{\text{tol}})^{-2}$ and $(\epsilon_{\text{tol}})^{-1}$.

3.2 The Singular Value Decomposition of A

Although we have now determined the condition number of A, it is possible to give a rather detailed analysis of its spectrum. This follows from the identification of A with the well-known prolate matrix, which was analyzed in detail by Slepian [31, 33]. We now review some of this work.

Following Slepian [31], let $P(N, W) \in \mathbb{C}^{N \times N}$ be the prolate matrix with entries

$$P(N, W)_{m,n} = \begin{cases} \frac{\sin 2\pi W(m-n)}{\pi(m-n)} & m \neq n, \\ 2W & m = n, \end{cases} \qquad m, n = 0, \ldots, N-1,$$

where $0 < W < \frac{1}{2}$ is fixed, and write $1 > \lambda_0(N, W) > \cdots > \lambda_{N-1}(N, W) > 0$ for its eigenvalues. Note that

$$\lambda_k\left(N, \frac{1}{2} - W\right) = 1 - \lambda_{N-1-k}(N, W). \tag{3.9}$$

The following asymptotic results are found in [31]:

(i) For fixed and small k,

$$1 - \lambda_k(N, W) \sim \sqrt{\pi}(k!)^{-1} 2^{(14k+9)/4} \alpha^{(2k+1)/4} (2-\alpha)^{-(k+1/2)} N^{k+1/2} \beta^{-N}, \tag{3.10}$$

where $\alpha = 1 - \cos 2\pi W$ and $\beta = \frac{\sqrt{2}+\sqrt{\alpha}}{\sqrt{2}-\sqrt{\alpha}}$.

(ii) For large N and k with $k = \lfloor 2WN(1-\epsilon) \rfloor$ and $0 < \epsilon < 1$, $1 - \lambda_k(N, W) \sim \mathrm{e}^{-c_1 - c_2 N}$ for explicitly known constants c_1, c_2 depending only on W and ϵ.

(iii) For large N and k with $k = \lfloor 2WN + (b/\pi) \log N \rfloor$, $\lambda_k(N, W) \sim \frac{1}{1+\mathrm{e}^{\pi b}}$.

(Slepian also derives similar asymptotic results for the eigenvectors of $P(N, W)$ [31].) From these results we conclude that the eigenvalues of the prolate matrix cluster exponentially near 0 and 1 and have a transition region of width $\mathcal{O}(\log N)$ around $k = 2WN$. This is shown in Fig. 1.

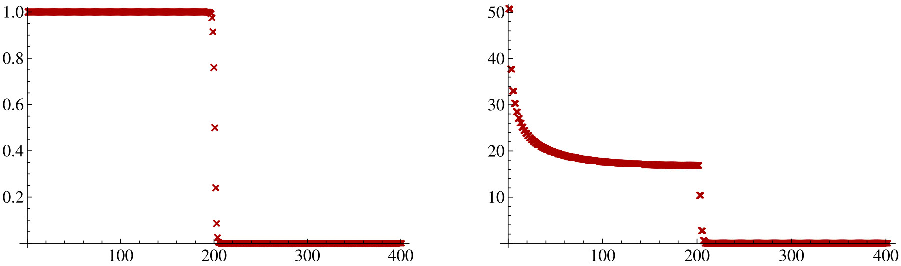


Fig. 1 Eigenvalues of the matrices (2.8) (*left*) and (2.13) (*right*) for $N = 200$ and $T = 2$

The matrix A of the continuous FE is precisely the prolate matrix $P(2N+1, \frac{1}{2T})$. In this case, the parameter β in (3.10) is given by

$$\beta = \frac{\sqrt{2}+\sqrt{\alpha}}{\sqrt{2}-\sqrt{\alpha}} = \cot^2\left(\frac{\pi}{4T}\right) = E(T).$$

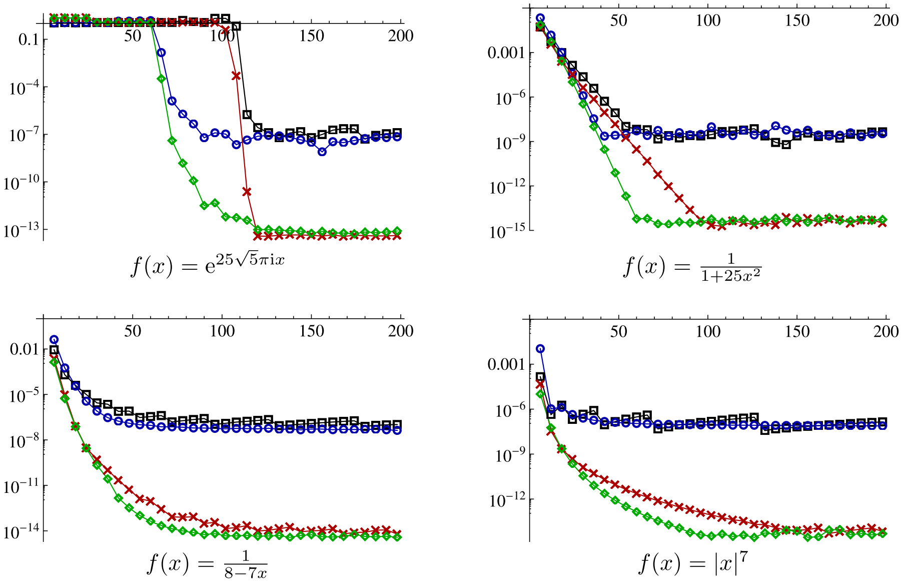
Applying Slepian's analysis, we now see that the eigenvalues of A cluster exponentially at rate $E(T)^2$ near zero and one (note that A corresponds to a prolate matrix of size $2N$), and in particular, that the condition number is $\mathcal{O}(E(T)^{2N})$. The latter estimate agrees with that given in Theorem 3.1. We remark, however, that Theorem 3.1 gives bounds for the minimal eigenvalue of A that hold for all N and T, unlike (3.10), which holds only for fixed T and sufficiently large N. Hence Theorem 3.1 remains valid when T is varied with N, an option which, as discussed in Sect. 2.4, can be advantageous in practice.

Since the matrix $\tilde{A}$ of the discrete FE is related to A (see Lemma 2.8), we expect a similar structure for its singular values. This is illustrated in Fig. 1. As is evident, the only qualitative difference between $\tilde{A}$ and A is found in the large singular values. The other features—the narrow transition region and the exponential clustering of singular values near 0—are much the same.

Remark 3.4 The choice $T = 2$ ($W = \frac{1}{4}$) is special. As shown by (3.9), the eigenvalues $\lambda_k(N, W)$ are symmetric in this case, and the transition region occurs at $k = \frac{1}{2}N$. This is unsurprising. When $T = 2$, the frame $\{e^{i\frac{n\pi}{2}x}\}_{n\in\mathbb{Z}}$ decomposes into two orthogonal bases, related to the sine and cosine transforms. Using this decomposition and the associated discrete transforms for each basis, M. Lyon has introduced an $\mathcal{O}(N(\log N)^2)$ complexity algorithm for computing FEs [25].

3.3 Numerical Examples

We now consider several numerical examples of the continuous and discrete FEs. In Fig. 2 we plot the error $\|f - f_N\|_\infty$ against N for various choices of f. Here the extension f_N is the numerically computed continuous or discrete FE—i.e. the result of solving the corresponding linear system in standard precision (recall Remark 1.2). Henceforth, we use the notation $G_N(f)$ and $\tilde{G}_N(f)$ for these *numerical* extensions, so as to distinguish them from their *exact* counterparts $F_N(f)$ and $\tilde{F}_N(f)$. Note that



$f(x) = \mathrm{e}^{25\sqrt{5}\pi \mathrm{i} x}$

$f(x) = \frac{1}{1+25x^2}$

$f(x) = \frac{1}{8-7x}$

$f(x) = |x|^7$

Fig. 2 The error $\|f - f_N\|_\infty$, where $f_N = G_N(f)$ (*squares* and *circles*) or $f_N = \tilde{G}_N(f)$ (*crosses* and *diamonds*) and $T = 2$ (*squares/crosses*) or $T = T(N; \epsilon_{\mathrm{tol}})$ (*circles/diamonds*) with $\epsilon_{\mathrm{tol}} = 10^{-14}$

the word ‘exact’ in this context refers to exact arithmetic. We do not mean exact in the sense that $F_N(f) = f$ for $f \in \mathcal{G}_N$.

At first sight, Fig. 2 appears somewhat surprising: for all three functions we obtain good accuracy, and there is no drift or growth in the error, even in the case where f is nonsmooth or has a complex singularity near $x = 0$. Evidently the ill-conditioning of the FE matrices established in Theorems 3.1 and 3.2 appears to have little effect on the numerical extensions $G_N(f)$ and $\tilde{G}_N(f)$. The purpose of Sect. 4 is to offer an explanation of this phenomenon.

In Fig. 2 we also compare two choices of T: fixed $T = 2$ and the N-dependent value (2.18) with $\epsilon_{\mathrm{tol}} = 10^{-14}$. Note that the latter typically outperforms the fixed value $T = 2$, especially for oscillatory functions. This is unsurprising in view of the discussion in Sect. 2.4.

Figure 2 also illustrates an important disadvantage of the continuous FE: namely, the approximation error levels off at around $\sqrt{\epsilon_{\mathrm{mach}}}$, where $\epsilon_{\mathrm{mach}} \approx 10^{-16}$ is the machine precision used, as opposed to around ϵ_{mach} for the discrete extension. Our analysis in Sect. 4 will confirm this phenomenon. Note that the differing behavior between the continuous and discrete extensions in this respect can be traced back to the observation made in Remark 3.3.

3.4 Condition Numbers of the Exact Continuous and Discrete FE Mappings

The exponential growth in the condition numbers of the continuous and discrete FE matrices imply extreme sensitivity in the FE coefficients to perturbations. However, the numerical results of Fig. 2 indicate that the FE approximations themselves are far more robust. Although we shall defer a full explanation of this difference to Sect. 4, it is possible to give a first insight by determining the condition numbers of the mappings F_N and $\tilde{F}_N$.

For vectors $b \in \mathbb{C}^{2N+1}$ and $\tilde{b} \in \mathbb{C}^{2N+2}$ let us write, with slight abuse of notation, $F_N(b)$ and $\tilde{F}_N(\tilde{b})$ for the corresponding continuous and discrete Fourier extensions whose coefficient vectors are the solutions of the linear systems $Aa = b$ and $\tilde{A}a = \tilde{b}$, respectively. We now define the condition numbers

$$\begin{aligned} \kappa(F_N) &= \sup\bigl\{ \bigl\| F_N(b) \bigr\| : b \in \mathbb{C}^{2N+1},\ \|b\| = 1 \bigr\}, \\ \kappa(\tilde{F}_N) &= \sup\bigl\{ \bigl\| F_N(b) \bigr\|_W : b \in \mathbb{C}^{2N+2},\ \|b\| = 1 \bigr\}. \end{aligned} \tag{3.11}$$

Here $\|\cdot\|$ denotes the usual l^2 vector norm, and W is the weight function of Lemma 2.8. Note that (3.11) gives the *absolute* condition numbers of F_N and $\tilde{F}_N$, as opposed to the more standard *relative* condition number [32]. The key results of this paper can easily be reformulated for the latter. However, we shall use (3.11) throughout, since it coincides with the definition given in [28] for linear mappings such as FEs. The work of [28] will be particularly relevant when considering equispaced FEs in Sect. 5.

We now have the following result:

Lemma 3.5 *The condition numbers of the exact continuous and discrete FEs satisfy*

$$\kappa(F_N) = 1/\sqrt{\lambda_{\min}(A)}, \qquad \kappa(\tilde{F}_N) = 1.$$

Proof Write $F_N(b) = \sum_{n=-N}^{N} a_n \phi_n$, where $Aa = b$. We have $\|F_N(b)\|^2 = a^* A a = b^* A^{-1} b$, and therefore $\kappa(F_N) = 1/\sqrt{\lambda_{\min}(A)}$, as required. For the second result, we note that $\|\tilde{F}_N(\tilde{b})\|^2 = a^* A_W a$, where $A_W = (\tilde{A})^* \tilde{A}$ is the matrix of Lemma 2.8. Since $\tilde{A}a = \tilde{b}$ the second result now follows. □

As with the FE matrices, this lemma shows that condition number of the discrete mapping $\tilde{F}_N$, which is identically one, is much better than that of the continuous mapping F_N. Similarly, the reason can be traced back to Remark 3.3. Note that this lemma establishes 2. of Sect. 1.

At first, it may seem that the fact that $\kappa(\tilde{F}_N) = 1$ explains the observed numerical stability in Fig. 2. However, since $\lambda_{\min}(A)$ is exponentially small (Theorem 3.1), the above lemma clearly does not explain the lack of drift in the numerical error in the case of the continuous FE. This is symptomatic of a larger issue: in general, the exact FEs $F_N(f)$ and $\tilde{F}_N(f)$ differ substantially from their numerical counterparts $G_N(f)$ and $\tilde{G}_N(f)$. As we show in the next section, there are important differences in both their stability *and* their convergence. In particular, any analysis based solely on F_N and $\tilde{F}_N$ is insufficient to describe the behavior of the numerical extensions G_N and $\tilde{G}_N$.

4 The Numerical Continuous and Discrete Fourier Extensions

We now analyze the numerical FEs G_N and $\tilde{G}_N$, and describe both when and how they differ from the exact extensions F_N and $\tilde{F}_N$.

4.1 The Norm of the Exact FE Coefficients

In short, the reason for this difference is as follows. Since the FE matrices A and $\tilde{A}$ are so ill-conditioned, the coefficients of the exact FEs F_N and $\tilde{F}_N$ will not usually be obtained in finite precision computations. To explain exactly how this affects stability and convergence, we first need to determine when this will occur. We require the following theorem:

Theorem 4.1 *Suppose that f is analytic in $\mathcal{D}(\rho^*)$ and continuous on its boundary. If $a \in \mathbb{C}^{2N+1}$ is the vector of coefficients of the continuous FE $F_N(f)$ then*

$$\|a\| \leq c_f \begin{cases} (\frac{E(T)}{\rho^*})^N & \rho^* < E(T), \\ N & \rho^* \geq E(T), \end{cases} \tag{4.1}$$

where c_f is proportional to $\max_{x \in \mathcal{D}(\rho)} |f(x)|$. *If $f \in \mathrm{L}^2(-1,1)$, then*

$$\|a\| \leq c \|f\| E(T)^N, \tag{4.2}$$

for some $c > 0$ independent of f and N.

Proof Write $F_N(f) = f_N = f_{\mathrm{e},N} + f_{\mathrm{o},N}$, where $f_{\mathrm{e},N}$ and $f_{\mathrm{o},N}$ are the even and odd parts of f_N, respectively. Since the set $\{\phi_n\}_{n\in\mathbb{Z}}$ is orthonormal over $[-T,T]$ we find that

$$\begin{aligned} \|a\| = \|f_N\|_{[-T,T]} &\leq 2\big(\|f_{\mathrm{e},N}\|_{[0,T]} + \|f_{\mathrm{o},N}\|_{[0,T]}\big) \\ &\leq 2\sqrt{T}\big(\|f_{\mathrm{e},N}\|_{\infty,[0,T]} + \|f_{\mathrm{o},N}\|_{\infty,[0,T]}\big). \end{aligned}$$

Recall from Sect. 2.1.1 that $f_{\mathrm{e},N}(x) = h_{1,N}(z)$ and $f_{\mathrm{o},N}(x) = \sin(\frac{\pi}{T} m^{-1}(z)) h_{2,N}(z)$, where $h_{i,N} \in \mathbb{P}_{N+1-i}$, $i = 1, 2$, is defined by (2.4). Thus, $\|a\| \leq c(\|h_{1,N}\|_{\infty,[m(T),1]} + \|h_{2,N}\|_{\infty,[m(T),1]})$ for some $c > 0$. Consider $h_{1,N}(z)$. This is precisely the expansion of the function $h_1(z) = f_1(m^{-1}(z))$ in polynomials $\{p_n\}_{n=0}^{\infty}$ orthogonal with respect to the weight function w_1: i.e. $h_{1,N} = \sum_{n=0}^{N} \langle h_1, p_n\rangle_{w_1} p_n$. Therefore

$$\|h_{1,N}\|_{\infty,[m(T),1]} \leq \sum_{n=0}^{N} \big|\langle h_1, p_n\rangle_{w_1}\big| \|p_n\|_{\infty,[m(T),1]}.$$

It is known that $\|p_n\|_{\infty,[m(T),1]} \leq c E(T)^n$ [22]. Also, since h_1 is analytic in $\mathcal{B}(\rho^*)$ we have $|\langle h_1, p_n\rangle_{w_1}| \leq c_f (\rho^*)^{-n}$. Hence

$$\|h_{1,N}\|_{\infty,[m(T),1]} \leq c_f \sum_{n=0}^{N} \big(E(T)/\rho^*\big)^n,$$

which gives (4.1). For (4.2) we use the bound $|\langle h_1, p_n\rangle_{w_1}| \leq \|h_1\|_{w_1} \leq c\|f\|$ instead. □

Corollary 4.2 *Let f be as in Theorem 4.1. Then the vector of coefficients $a \in \mathbb{C}^{2N+2}$ of the discrete Fourier extension $\tilde{F}_N(f)$ of f satisfies the same bounds as those given in Theorem 4.1.*

Proof The functions $h_{i,N}$, $i = 1, 2$ are the polynomial interpolants of h_i at the nodes (2.10) (Proposition 2.7). Write $h_{i,N}(z) = \sum_{n=0}^{N} \tilde{d}_n T_n(z)$, where $T_n(z)$ is the nth Chebyshev polynomial, and let $\hat{d}_n = \langle h_i, T_n\rangle_w$ be the Chebyshev polynomial coefficient of h_i. Note that $|\hat{d}_n| \leq c_f(\rho^*)^{-n}$. Due to aliasing formula $\tilde{d}_n = \hat{d}_n + \sum_{k\neq 0}(\hat{d}_{2kN+n} + \hat{d}_{2kN-n})$ (see [13, Eq. (2.4.20)]) we obtain

$$|\tilde{d}_n| \leq c_f\left((\rho^*)^{-n} + \sum_{k=1}^{\infty}(\rho^*)^{-2kN-n} + \sum_{k=1}^{\infty}(\rho^*)^{-2kN+n}\right)$$
$$\leq c_f\big((\rho^*)^{-n} + (\rho^*)^{n-2N}\big) \leq c_f(\rho^*)^{-n}.$$

The result now follows along the same lines as the proof of Theorem 4.1. □

To compute the continuous or discrete FE we need to solve the linear system $Aa = b$ (respectively $\tilde{A}a = \tilde{b}$). When N is large, the columns of A ($\tilde{A}$) become near-linearly dependent, and, as shown in Sect. 3.2, the numerical rank of A is roughly $1/T$ times its dimension. Now suppose we solve $Aa = b$ with a standard numerical solver. Loosely speaking, the solver will use the extra degrees of freedom to construct approximate solutions a' with small norms. The previous theorem and corollary therefore suggest the following. In general, only in those cases where f is analytic with $\rho^* \geq E(T)$ can we expect the theoretical coefficient vector a to be produced by the numerical solver for all N. Outside of this case, we may well have $a' \neq a$ for sufficiently large N, due to the potential for exponential growth of the latter. Hence, in this case, the numerical extension $G_N(f)$ will not coincide with the exact extension $F_N(f)$.

This raises the following question: if the numerical solver does not give the exact coefficients vector, then what does it yield? The following proposition confirms the existence of infinitely many approximate solutions of the equations $Aa = b$ with small norm coefficient vectors:

Proposition 4.3 *Suppose that $f \in \mathrm{H}^k(-1,1)$. Then there exist $a^{[N]} \in \mathbb{C}^{2N+1}$, $N \in \mathbb{N}$, satisfying*

$$\big\|a^{[N]}\big\| \leq c_k(T)\|f\|_{\mathrm{H}^k(-1,1)}, \tag{4.3}$$

and

$$\big\|Aa^{[N]} - b\big\| \leq c_k(T)N^{-k}\|f\|_{\mathrm{H}^k(-1,1)}, \tag{4.4}$$

where $c_k(T)$ is the constant of Lemma 2.5. Moreover, if $g_N = \sum_{|n|\leq N} a_n^{[N]}\phi_n$ then

$$\|f - g_N\| \leq c_k(T)N^{-k}\|f\|_{\mathrm{H}^k(-1,1)}. \tag{4.5}$$

Proof Let $\tilde{f} \in \mathrm{H}^k(\mathbb{T})$ be the extension guaranteed by Lemma 2.5, and write $a^{[N]}$ for the vector of its first $2N+1$ Fourier coefficients on $\mathbb{T} = [-T, T)$. By Bessel's inequality, $\|a^{[N]}\| \leq \|\tilde{f}\|_{[-T,T]} \leq c_k(T)\|f\|_{\mathrm{H}^k(-1,1)}$, which gives (4.3). For (4.4), we merely note that $(Aa^{[N]} - b)_n = \langle f - g_N, \phi_n \rangle$. Using the frame property (2.5) we obtain $\|Aa^{[N]} - b\| \leq \|f - g_N\|$. Thus, (4.4) follows directly from (4.5), and the latter is a standard result of Fourier analysis (see [13, eq. (5.1.10)], for example). □

This proposition states that there exist vectors with norms bounded independently of N that solve the equations $Aa = b$ up to an error of order N^{-k}. Moreover, these vectors yield extensions which converge algebraically fast to f at rate k. Whilst it does not imply that these are the vectors produced by the numerical solver, it does indicate that, in the case where the exact extension $F_N(f)$ or $\tilde{F}_N(f)$ has a large coefficient norm, geometric convergence of the numerical extension $G_N(f)$ or $\tilde{G}_N(f)$ may be sacrificed for superalgebraic convergence so as to retain boundedness of the computed coefficients.

This hypothesis is verified numerically in Fig. 3 (all computations were carried out in *Mathematica*, with additional precision used to compute the exact FEs and standard precision used otherwise). Geometric convergence of the exact extension is replaced by slower, but still high-order convergence for sufficiently large N. Note that the 'breakpoint' occurs at roughly the same value of N regardless of the function being approximated. Moreover, the breakpoint occurs at a larger value of N for the discrete extension than for the continuous extension.

These observations will be established rigorously in the next section. However, we now make several further comments on Fig. 3. First, note that the breakdown of geometric convergence is far less severe for the classical Runge function $f(x) = \frac{1}{1+16x^2}$ than for the functions $f(x) = \frac{1}{8-7x}$ and $f(x) = 1 + \frac{\cosh 40x}{\cosh 40}$. This can be explained by the behavior of these functions near $x = \pm 1$. The Runge function $f(x) = \frac{1}{1+16x^2}$ is reasonably flat near $x = \pm 1$. Hence it possesses extensions with high degrees of smoothness which do not grow large on the extended domain $[-T, T]$. Conversely, the other two functions have boundary layers near $x = 1$ (also $x = -1$ for the latter). Therefore any smooth extension will be large on $[-T, T]$, and by Parseval's relation, the coefficient vectors corresponding to the Fourier series of this extension will also have large norm.

Second, although it is not apparent from Fig. 3 that the convergence rate beyond the breakpoint is truly superalgebraic, this is in fact the case. This is confirmed by Fig. 4. In the right-hand diagram we plot the error against N in log–log scale. The slight downward curve in the error indicates superalgebraic convergence. Had the convergence rate been algebraic of fixed order, then the error would have followed a straight line.

4.2 Analysis of the Numerical Continuous and Discrete FEs

We now wish to analyze the numerical extensions $G_N(f)$ and $\tilde{G}_N(f)$. Since the numerical solvers used in environments such as *Matlab* or *Mathematica* are difficult to analyze directly, we shall look at the result of solving $Aa = b$ (or $\tilde{A}a = \tilde{b}$) with a truncated singular value decomposition (SVD). This represents an idealization of

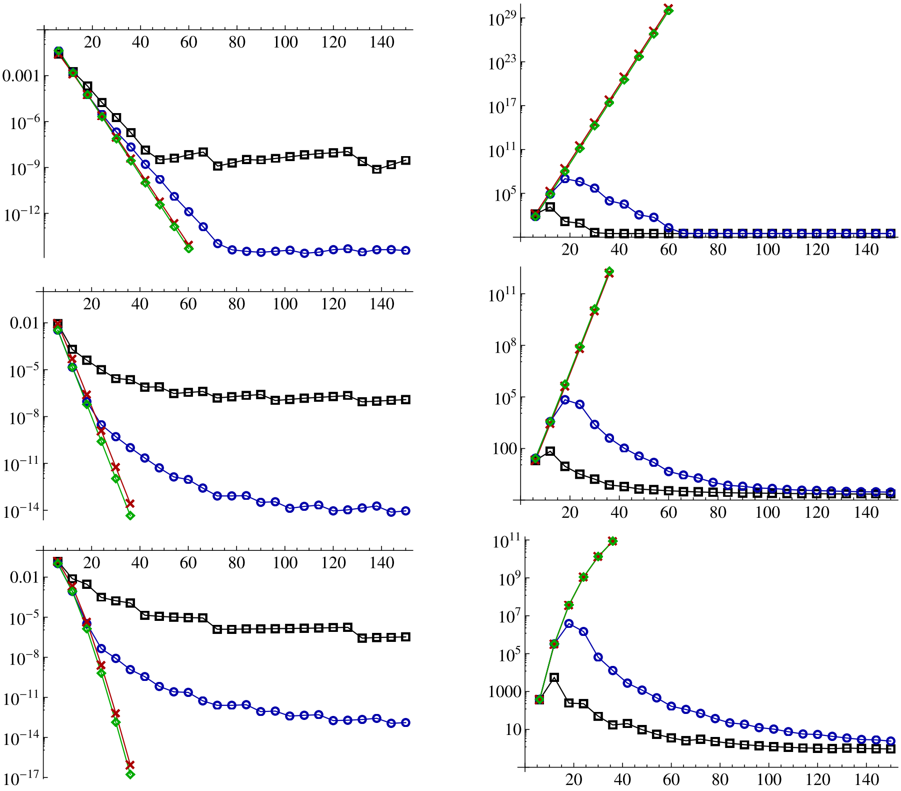


Fig. 3 Comparison of the numerical continuous and discrete FEs $G_N(f)$ and $\tilde{G}_N(f)$ (*squares* and *circles*) and their exact counterparts $F_N(f)$ and $\tilde{F}_N(f)$ (*crosses* and *diamonds*) for $T = 2$. *Left*: the uniform error $\|f - f_N\|_\infty$ against N. *Right*: the norm $\|a\|$ of the coefficient vector. *Top row*: $f(x) = \frac{1}{1+16x^2}$. *Middle row*: $f(x) = \frac{1}{8-7x}$. *Bottom row*: $f(x) = 1 + \frac{\cosh 40x}{\cosh 40}$

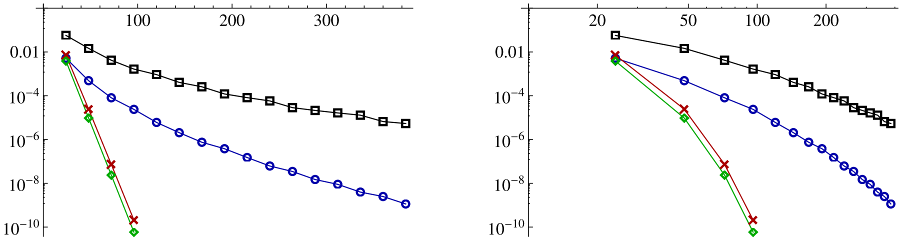


Fig. 4 Comparison of the numerical continuous and discrete FEs $G_N(f)$ and $\tilde{G}_N(f)$ (*squares* and *circles*) and their exact counterparts $F_N(f)$ and $\tilde{F}_N(f)$ (*crosses* and *diamonds*) for $T = 2$ and $f(x) = \frac{1}{101-100x}$. *Left*: the uniform error in log scale. *Right*: the uniform error in log–log scale

the numerical solver. Indeed, neither *Matlab*'s \ or *Mathematica*'s `LeastSquares` actually performs a truncated SVD. However, in practice, this simplification appears reasonable: numerical experiments indicate that these standard solvers give roughly the same approximation errors as the truncated SVD with suitably small truncation parameter (typically $\epsilon = 10^{-14}$). We shall also assume throughout that the truncated SVD is computed without error. However, this also seems fair: in experiments, we observe that the finite-precision SVD gives similar results to the numerical solver whenever the tolerance is sufficiently small.

Suppose that A (respectively $\tilde{A}$) has SVD USV^* with S being the diagonal matrix of singular values. Given a truncation parameter $\epsilon > 0$, we now consider the solution

$$a_\epsilon = VS^\dagger U^* b, \tag{4.6}$$

where $S^\dagger$ is the diagonal matrix with nth entry $1/\sigma_n$ if $\sigma_n > \epsilon$ and 0 otherwise. We write

$$H_{N,\epsilon}(f) = \sum_{|n|\leq N} (a_\epsilon)_n \phi_n,$$

for the corresponding FE. Suppose that $v_n \in \mathbb{C}^{2N+1}$ is the right singular vector of A with singular value σ_n, and let

$$\Phi_n = \sum_{|m|\leq N} (v_n)_m \phi_m \in \mathcal{G}_N,$$

be the Fourier series corresponding to v_n. Note that the functions Φ_n are orthonormal with respect to $\langle \cdot, \cdot \rangle_{[-T,T]}$ and span $\mathcal{G}_N$. Also, if we define $\mathcal{G}_{N,\epsilon} = \operatorname{span}\{\Phi_n : \sigma_n > \epsilon\} \subseteq \mathcal{G}_N$, then we have $H_{N,\epsilon}(f) \in \mathcal{G}_{N,\epsilon}$.

We now consider the cases of the continuous and discrete FEs separately.

4.2.1 The Continuous Fourier Extension

In this case, since A is Hermitian and positive definite, the singular vectors v_n are actually eigenvectors of A with $Av_n = \sigma_n v_n$. By definition, we have $\langle \Phi_n, \Phi_m \rangle = (v_n)^* A v_m = \sigma_n \delta_{n,m}$, and therefore

$$H_{N,\epsilon}(f) = \sum_{n:\sigma_n > \epsilon} \frac{1}{\sigma_n} \langle f, \Phi_n \rangle \Phi_n. \tag{4.7}$$

Our main result is as follows:

Theorem 4.4 *Let $f \in \mathrm{L}^2(-1,1)$ and suppose that $H_{N,\epsilon}(f)$ is given by (4.7). Then*

$$\left\| f - H_{N,\epsilon}(f) \right\| \leq \| f - \phi \| + \sqrt{\epsilon} \| \phi \|_{[-T,T]}, \quad \forall \phi \in \mathcal{G}_N, \tag{4.8}$$

and

$$\| a_\epsilon \| = \left\| H_{N,\epsilon}(f) \right\|_{[-T,T]} \leq \frac{1}{\sqrt{\epsilon}} \| f - \phi \| + \| \phi \|_{[-T,T]}, \quad \forall \phi \in \mathcal{G}_N. \tag{4.9}$$

Proof The function $H_{N,\epsilon}(f)$ is the orthogonal projection of f onto $\mathcal{G}_{N,\epsilon}$ with respect to $\langle\cdot,\cdot\rangle$. Hence for any $\phi\in\mathcal{G}_N$ we have $\|f-H_{N,\epsilon}(f)\|\le\|f-H_{N,\epsilon}(\phi)\|\le\|f-\phi\|+\|\phi-H_{N,\epsilon}(\phi)\|$. Consider the latter term. Since $\phi\in\mathcal{G}_N$, the observation that the functions Φ_n are also orthonormal on $[-T,T]$ gives

$$\begin{aligned}\big\|\phi-H_{N,\epsilon}(\phi)\big\|^2&=\Bigg\|\sum_{n:\sigma_n<\epsilon}\langle\phi,\Phi_n\rangle_{[-T,T]}\Phi_n\Bigg\|^2\\&=\sum_{n:\sigma_n<\epsilon}\sigma_n\big|\langle\phi,\Phi_n\rangle_{[-T,T]}\big|^2\le\epsilon\|\phi\|^2_{[-T,T]}.\end{aligned}$$

This yields (4.8). For (4.9) we first write $\|H_{N,\epsilon}(f)\|_{[-T,T]}\le\|H_{N,\epsilon}(f-\phi)\|_{[-T,T]}+\|H_{N,\epsilon}(\phi)\|_{[-T,T]}$. By orthogonality,

$$\begin{aligned}\big\|H_{N,\epsilon}(f-\phi)\big\|^2_{[-T,T]}&=\sum_{n:\sigma_n>\epsilon}\frac{1}{\sigma_n^2}\big|\langle f-\phi,\Phi_n\rangle\big|^2\le\frac{1}{\epsilon}\sum_{n:\sigma_n>\epsilon}\frac{1}{\sigma_n}\big|\langle f-\phi,\Phi_n\rangle\big|^2\\&=\frac{1}{\epsilon}\big\|H_{N,\epsilon}(f-\phi)\big\|^2.\end{aligned}$$

Since $H_{N,\epsilon}$ is an orthogonal projection, we conclude that $\|H_{N,\epsilon}(f-\phi)\|^2_{[-T,T]}\le\frac{1}{\epsilon}\|f-\phi\|^2$, which gives the first term in (4.9). For the second, we notice that

$$\big\|H_{N,\epsilon}(\phi)\big\|^2_{[-T,T]}=\sum_{n:\sigma_n>\epsilon}\big|\langle\phi,\Phi_n\rangle_{[-T,T]}\big|^2\le\|\phi\|^2_{[-T,T]},$$

since $\phi\in\mathcal{G}_N$. □

This theorem allows us to explain the behavior of the numerical FE $G_N(f)$. Suppose that f is analytic in $\mathcal{D}(\rho)$ and continuous on its boundary, where $\rho<E(T)$ and $\mathcal{D}(\rho)$ is as in Theorem 2.11. Set $\phi=F_N(f)$ in (4.8), where $F_N(f)$ is the exact continuous FE. Then Theorems 2.11 and 4.1 give

$$\big\|f-H_{N,\epsilon}(f)\big\|\le c_f\big(1+\sqrt{\epsilon}E(T)^N\big)\rho^{-N}. \tag{4.10}$$

For small N, the first term in the brackets dominates, and we see geometric convergence of $H_{N,\epsilon}(f)$, and therefore also $G_N(f)$, at rate ρ. Convergence continues as such until the breakpoint

$$N_0=N_0(\epsilon,T):=-\frac{\log\epsilon}{2\log E(T)}, \tag{4.11}$$

at which point the second term dominates and the bound begins to increase. On the other hand, Proposition 4.3 establishes the existence of functions $\phi\in\mathcal{G}_N$ with bounded coefficients which approximate f to any given algebraic order. Substituting such a function ϕ into (4.8) gives

$$\big\|f-H_{N,\epsilon}(f)\big\|\le c_k(T)\big(N^{-k}+\sqrt{\epsilon}\big)\|f\|_{\mathrm{H}^k(-1,1)},\quad\forall N,k\in\mathbb{N}. \tag{4.12}$$

Therefore, once $N > N_0(\epsilon, T)$ we expect at least superalgebraic convergence of $H_{N,\epsilon}(f)$ down to a maximal achievable accuracy of order $\sqrt{\epsilon}\|f\|$. Note that at the breakpoint $N = N_0$, the error satisfies

$$\|f - H_{N_0,\epsilon}(f)\| \leq 2c_f\big(\sqrt{\epsilon}\big)^{d_f}, \quad d_f = \frac{\log\rho}{\log E(T)} \in (0, 1]. \tag{4.13}$$

If f is analytic in $\mathcal{D}(E(T))$, and if $c_f = \max_{x\in\mathcal{D}(\rho)}|f(x)|$ is not too large, then f is already approximated to order $\sqrt{\epsilon}$ accuracy at this point. It is only in those cases where either $\rho < E(T)$ or where c_f is large (or both) that one sees the second phase of superalgebraic convergence.

Theorem 4.1 also explains the behavior of the coefficient norm $\|a_\epsilon\|$. Observe that breakpoint $N_0(\epsilon, T)$ is (up to a small constant) the largest N for which all singular values of A are included in its truncated SVD (see Theorem 3.1). Thus, when $N < N_0(\epsilon, T)$, we have $H_{N,\epsilon}(f) = F_N(f)$, and Theorem 4.1 indicates exponential growth of $\|a_\epsilon\|$. On the other hand, once $N > N_0(\epsilon, T)$, we use (4.9) to obtain

$$\|a_\epsilon\| \leq c_k(T)\big(N^{-k}/\sqrt{\epsilon} + 1\big)\|f\|_{\mathrm{H}^k(-1,1)}, \quad \forall N, k \in \mathbb{N}.$$

In particular, for $N > N_0(\epsilon, T)$, we expect decay of $\|a_\epsilon\|$ down from its maximal value at $N = N_0(\epsilon, T)$.

This analysis is corroborated in Fig. 5, where we plot the error and coefficient norm for the truncated SVD extension for various test functions. Note that the maximal achievable accuracy in all cases is order $\sqrt{\epsilon}$, consistently with our analysis. Moreover, for the meromorphic functions $f(x) = \frac{1}{1+16x^2}$ and $f(x) = \frac{1}{8-7x}$ we see initial geometric convergence followed by slower convergence after N_0, again as our analysis predicts. The qualitative difference in convergence for these functions in the regime $N > N_0$ is due to the contrasting behavior of their derivatives (recall the discussion in Sect. 4.1). On the other hand, the convergence effectively stops at N_0 for $f(x) = x$, since this function has small constant c_f and is therefore already resolved down to order $\sqrt{\epsilon}$ when $N = N_0$.

Since $N_0(10^{-6}, 2) \approx 4$, $N_0(10^{-12}, 2) \approx 8$, $N_0(10^{-18}, 2) \approx 12$, and $N_0(10^{-24}, 2) \approx 16$, Fig. 5 also confirms the expression (4.11) for the breakpoint in convergence. In particular, the breakpoint is independent of the function being approximated. This latter observation is unsurprising. As noted, $N_0(\epsilon, T)$ is the largest value of N for which $H_{N,\epsilon}(f)$ coincides with $F_N(f)$. Beyond this point, $H_{N,\epsilon}(f)$ will not typically agree with $F_N(f)$, and thus we cannot expect further geometric convergence in general. Note that our analysis does not rule out geometric convergence for $N > N_0$. There may well be certain functions for which this occurs. However, extensive numerical tests suggest that in most cases, one sees only superalgebraic convergence in this regime, and indeed, this is all that we have proved.

Remark 4.5 At first sight, it may appear counterintuitive that one can still obtain good accuracy when excluding all singular values below a certain tolerance. However, recall that we are not interested in the accuracy of computing a, but rather the accuracy of $F_N(f)$ on the domain $[-1, 1]$. Since the nth singular value σ_n is equal

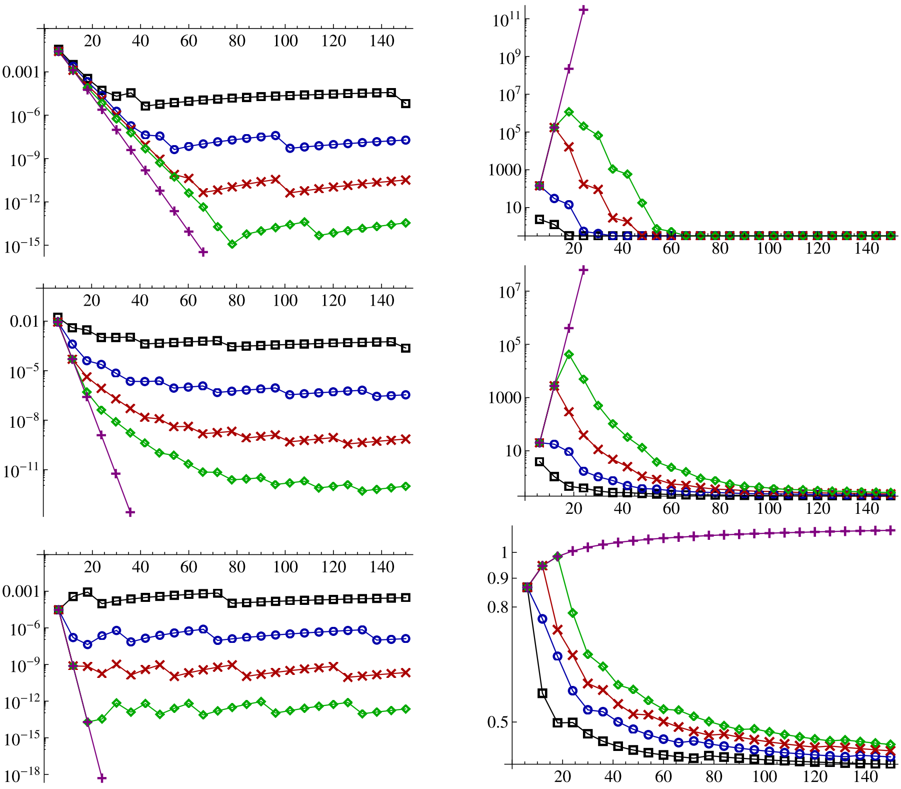


Fig. 5 Error (*left*) and coefficient norm (*right*) against N for the continuous FE with $T=2$, where $f(x)=\frac{1}{1+16x^2}$ (*top row*), $f(x)=\frac{1}{8-7x}$ (*middle row*) and $f(x)=x$ (*bottom row*). *Squares*, *circles*, *crosses* and *diamonds* correspond to the truncated SVD extension $H_{N,\epsilon}(f)$ with $\epsilon=10^{-6}, 10^{-12}, 10^{-18}, 10^{-24}$, respectively, and pluses correspond to the exact extension $F_N(f)$

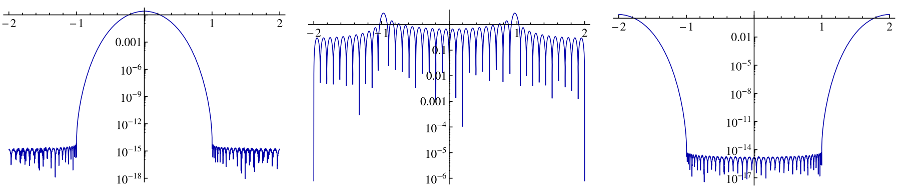


Fig. 6 The SVD functions $|\Phi_n(x)|$ for $n=0$, $n=20$ and $n=40$, where $N=20$ and $T=2$

to $\|\Phi_n\|^2/\|\Phi_n\|^2_{[-T,T]}$, the functions Φ_n excluded from $H_{N,\epsilon}(f)$ are precisely those for which $\|\Phi_n\|^2<\epsilon\|\Phi_n\|^2_{[-T,T]}$. In other words, they have little effect on $F_N(f)$ in $[-1,1]$.

In Fig. 6 we plot the functions Φ_n for several n. Note that these functions are precisely the discrete prolate spheroidal wavefunctions of Slepian [31]. As predicted, when n is small, the function Φ_n is large in $[-1,1]$ and small in $[-T,T]\backslash[-1,1]$.

When n is in the transition region ($n \approx 2N/T$, see Sect. 3.2), the function Φ_n is roughly of equal magnitude in both regions, and for $n \approx 2N$, Φ_n is much smaller in $[-1,1]$ than on $[-T,T]$. Note also that Φ_n is increasingly oscillatory in $[-1,1]$ as n increases, and decreasingly oscillatory in $[-T,T]\backslash[-1,1]$. This follows from the fact that Φ_n has precisely n zeroes in $[-1,1]$ and $2N-n$ zeroes in $[-T,T]\backslash[-1,1]$ [31]. Such behavior also implies that any 'nice' function will eventually be well approximated by functions Φ_n corresponding to 'nice' eigenvalues, as expected.

4.2.2 The Discrete Fourier Extension

In this case, we have $(\Phi_n,\Phi_m)_N = \sigma_n^2 \delta_{n,m}$, where

$$(f,g)_N = \frac{\pi}{N+1} \sum_{n=-N-1}^{N} f(x_n)\overline{g(x_n)},$$

is the discrete inner product corresponding to the quadrature nodes $\{x_n\}_{n=-N-1}^{N}$. Therefore

$$\tilde{H}_{N,\epsilon}(f) = \sum_{n:\sigma_n>\epsilon} \frac{1}{\sigma_n^2}(f,\Phi_n)_N \Phi_n \in \mathcal{G}'_{N,\epsilon} := \operatorname{span}\{\Phi_n : \sigma_n > \epsilon\}, \tag{4.14}$$

is the orthogonal projection of f onto $\mathcal{G}'_{N,\epsilon}$ with respect to the discrete inner product $(\cdot,\cdot)_N$.

Theorem 4.6 *Let $f \in \mathrm{L}^\infty(-1,1)$ and $\tilde{H}_{N,\epsilon}(f)$ be given by (4.14). Then*

$$\left\|f-\tilde{H}_{N,\epsilon}(f)\right\|_W \le \|f-\phi\|_W + \sqrt{2\pi Q(N;\epsilon)}\|f-\phi\|_\infty + \epsilon\|\phi\|_{[-T,T]}, \quad \forall \phi \in \mathcal{G}_N, \tag{4.15}$$

and

$$\|a_\epsilon\| = \left\|\tilde{H}_{N,\epsilon}(f)\right\|_{[-T,T]} \le \frac{1}{\epsilon}\sqrt{2\pi Q(N;\epsilon)}\|f-\phi\|_\infty + \|\phi\|_{[-T,T]}, \quad \forall \phi \in \mathcal{G}_N, \tag{4.16}$$

where $Q(N;\epsilon) = |\{n : \sigma_n > \epsilon\}| \le 2(N+1)$ and W is the weight function of Lemma 2.8.

Proof By the triangle inequality,

$$\left\|f-\tilde{H}_{N,\epsilon}(f)\right\|_W \le \|f-\phi\|_W + \left\|\phi-\tilde{H}_{N,\epsilon}(\phi)\right\|_W + \left\|\tilde{H}_{N,\epsilon}(f-\phi)\right\|_W, \quad \forall \phi \in \mathcal{G}'_N.$$

Consider the second term. Since $\phi \in \mathcal{G}'_N$, and the quadrature is exact on $\mathcal{G}'_N$, we have

$$\begin{aligned}\left\|\phi-\tilde{H}_{N,\epsilon}(\phi)\right\|_W^2 &= \big(\phi-\tilde{H}_{N,\epsilon}(\phi), \phi-\tilde{H}_{N,\epsilon}(\phi)\big)_N \\ &= \sum_{n:\sigma_n<\epsilon} \sigma_n^2 \left|\langle \phi,\Phi_n\rangle_{[-T,T]}\right|^2 \le \epsilon^2 \|\phi\|_{[-T,T]}^2.\end{aligned}$$

For the third term, let g be arbitrary. Then $(\tilde{H}_{N,\epsilon}(g), \tilde{H}_{N,\epsilon}(g))_N = \sum_{n:\sigma_n>\epsilon} \frac{1}{\sigma_n^2}|(g, \Phi_n)_N|^2$. Hence

$$\| \tilde{H}_{N,\epsilon}(g) \|_W^2 = \big(\tilde{H}_{N,\epsilon}(g), \tilde{H}_{N,\epsilon}(g)\big)_N \leq (g, g)_N \sum_{n:\sigma_n>\epsilon} \frac{1}{\sigma_n^2} (\Phi_n, \Phi_n)_N$$
$$= (g, g)_N Q(N; \epsilon), \tag{4.17}$$

since $(\Phi_n, \Phi_n)_N = \sigma_n^2$. It is straightforward to show that $(g, g)_N \leq 2\pi \|g\|_\infty^2$. Setting $g = f - \phi$ now gives the corresponding term in (4.15), and completes its proof. For (4.16), we proceed as in the proof of Theorem 4.4. Note that

$$\| \tilde{H}_{N,\epsilon}(g) \|_{[-T,T]}^2 = \sum_{n:\sigma_n>\epsilon} \frac{1}{\sigma_n^4} \big|(g, \Phi_n)_N\big|^2 \leq \frac{1}{\epsilon^2} \| \tilde{H}_{N,\epsilon}(g) \|_W^2, \tag{4.18}$$

for any $g \in \mathrm{L}^\infty(-1, 1)$. Also,

$$\big\| \tilde{H}_{N,\epsilon}(\phi) \big\|_{[-T,T]} \leq \|\phi\|_{[-T,T]}, \quad \phi \in \mathcal{G}_N. \tag{4.19}$$

The result now follows by writing $\|\tilde{H}_{N,\epsilon}(f)\|_{[-T,T]} \leq \|\tilde{H}_{N,\epsilon}(f - \phi)\|_{[-T,T]} + \|\tilde{H}_{N,\epsilon}(\phi)\|_{[-T,T]}$ and using (4.17)–(4.19). □

As with the continuous FE, this theorem allows us to analyze the numerical discrete extension $\tilde{G}_N(f)$. Once more we deduce geometric convergence in N up to the function-independent breakpoint

$$N_1(T; \epsilon) := -\frac{\log \epsilon}{\log E(T)} \equiv 2N_0(T; \epsilon), \tag{4.20}$$

with superalgebraic convergence beyond this point. These conclusions are confirmed in Fig. 7. Note, however, two key differences between the continuous and discrete FE. First, the bound (4.15) involves ϵ, as opposed to $\sqrt{\epsilon}$, meaning that we expect convergence of $\tilde{G}_N(f)$ down to close to machine precision. Second, the breakpoint $N_1(T; \epsilon)$ is precisely twice $N_0(T; \epsilon)$. Hence, the regime of geometric convergence of $\tilde{G}_N(f)$ is exactly twice as large as that of the continuous FE. These observations are in close agreement with the behavior seen in the numerical examples in Sect. 4.1.

4.3 The Condition Numbers of the Numerical Continuous and Discrete FEs

Having analyzed the convergence of the numerical FE—and in particular, established 5. of Sect. 1—we next address its condition number. Once more, we do this by considering the extensions $H_{N,\epsilon}$ and $\tilde{H}_{N,\epsilon}$:

Theorem 4.7 *Let $H_{N,\epsilon}$ be the continuous truncated SVD FE given by* (4.7). *Then*

$$\kappa(H_{N,\epsilon}) = 1/\min\big\{\sqrt{\sigma_n} : \sigma_n > \epsilon\big\} \leq \min\big\{1/\sqrt{\epsilon}, c(T)N^{\frac{3}{2}}E(T)^N\big\}, \quad N \in \mathbb{N}, \ \epsilon > 0,$$

where $c(T)$ is a positive constant independent of N. Conversely, if $\tilde{H}_{N,\epsilon}$ is the discrete extension (4.14), *then $\kappa(\tilde{H}_{N,\epsilon}) = 1$ for all $N \in \mathbb{N}$ and $\epsilon > 0$.*

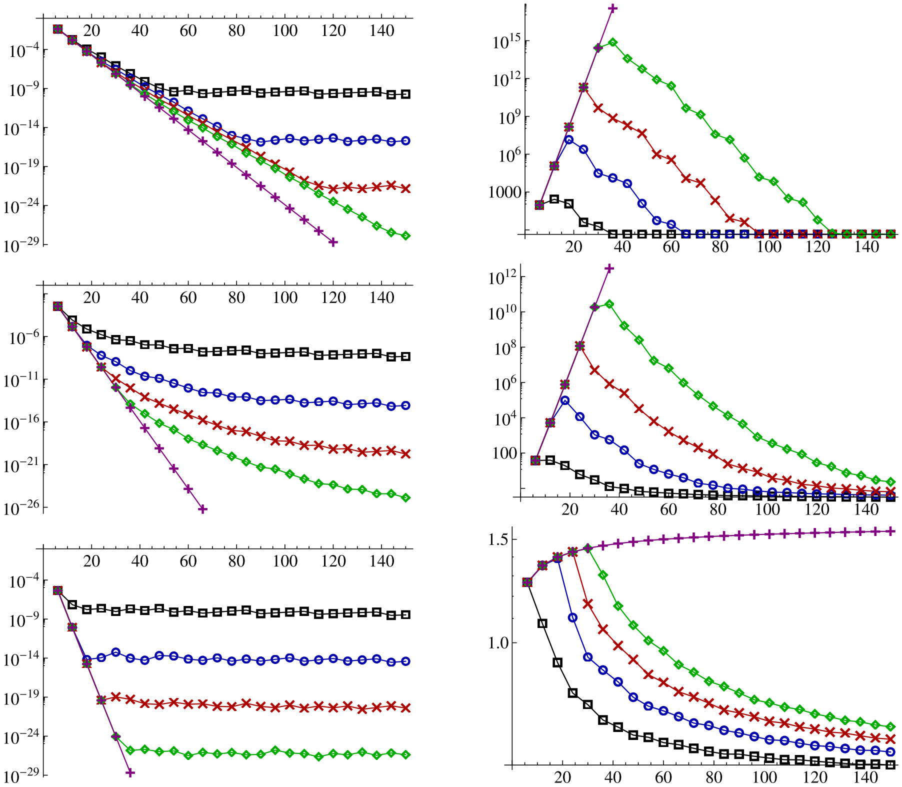


Fig. 7 Error (*left*) and coefficient norm (*right*) against N for the discrete FE with $T = 2$, where $f(x) = \frac{1}{1+16x^2}$ (*top row*), $f(x) = \frac{1}{8-7x}$ (*middle row*) and $f(x) = x$ (*bottom row*). *Squares*, *circles*, *crosses* and *diamonds* correspond to the truncated SVD extension $H_{N,\epsilon}(f)$ with $\epsilon = 10^{-6}, 10^{-12}, 10^{-18}, 10^{-24}$, respectively, and pluses correspond to the exact extension $F_N(f)$

Proof The proof of the equalities is similar to that of Lemma 3.5 with A and $\tilde{A}$ replaced by their truncated SVD versions. The upper bound for $\kappa(H_{N,\epsilon})$ is a consequence of Theorem 3.1. □

This theorem, which establishes 3. of Sect. 1, has some interesting consequences. First, the discrete FE is perfectly stable. On the other hand, the numerical continuous FE is far from stable. The condition number grows exponentially fast at rate $E(T)$ until it reaches $1/\sqrt{\epsilon}$, where ϵ is the truncation parameter in the SVD. Thus, with the continuous FE, we may see perturbations being magnified by a factor of $1/\sqrt{\epsilon_{\text{mach}}} \approx 10^8$ in practice.

Note that G_N and $\tilde{G}_N$ are both substantially better conditioned than the corresponding coefficient mappings. The explanation for this difference comes from Remark 4.5. A perturbation η in the input vector b gives large errors in the FE coefficients if η has a significant component in the direction of a singular vector v_n associated with a small singular value σ_n. However, since the corresponding function Φ_n

Table 1 The functions $K(G_N)$ and $K(\tilde{G}_N)$ for $T=2$

N	40	80	120	160	200
$K(G_N)$	4.93×10^6	4.22×10^6	3.30×10^6	3.82×10^6	5.28×10^6
$K(\tilde{G}_N)$	8.00×10^0	1.04×10^1	1.23×10^1	1.39×10^1	1.53×10^1

is small on $[-1,1]$, this error is substantially reduced (in the case of the continuous FE) or canceled out altogether (for the discrete FE) in the resulting extension.

Another implication of Theorem 4.7 is the following: varying T has no substantial effect on stability. Although the condition number of the FE matrices depends on T (recall Theorems 3.1 and 3.2), as does the condition number of the exact continuous FE (see Lemma 3.5), the condition numbers of the numerical mappings $\tilde{G}_N$ and, for all large N, G_N are actually independent of this parameter.

It is important to confirm that the results of this theorem on the condition number of the truncated SVD extensions predict the behavior of the numerical extensions G_N and $\tilde{G}_N$. It is easiest to do this by computing upper bounds for $\kappa(G_N)$ and $\kappa(\tilde{G}_N)$. Let $\{e_n\}_{n=1}^{2N+1}$ be the standard basis for $\mathbb{C}^{2N+1}$. Then a simple argument gives

$$\left\|G_N(b)\right\| \leq \|b\| \sqrt{\sum_{n=1}^{2N+1} \left\|G_N(e_n)\right\|^2}, \quad \forall b \in \mathbb{C}^{2N+1}, \tag{4.21}$$

and therefore

$$\kappa(G_N) \leq K(G_N) := \sqrt{\sum_{n=1}^{2N+1} \left\|G_N(e_n)\right\|^2}. \tag{4.22}$$

We define the upper bound $K(\tilde{G}_N)$ in a similar manner:

$$\kappa(\tilde{G}_N) \leq K(\tilde{G}_N) := \sqrt{\sum_{n=1}^{2N+2} \left\|\tilde{G}_N(e_n)\right\|_W^2}.$$

In Table 1 we show $K(G_N)$ and $K(\tilde{G}_N)$ for various choices of N. As we see, the discrete FE is extremely stable: not only is there no blowup in N, but the value of $K(\tilde{G}_N)$ is also close to one in magnitude. For the continuous extension, we see that $K(G_N) \approx 5 \times 10^6 = 1/\sqrt{\epsilon}$, where $\epsilon = 2.5 \times 10^{-13}$. This behavior is in good agreement with Theorem 4.7.

The difference in stability between the continuous and discrete FEs is highlighted in Fig. 8. Here we perturbed the right-hand side b of the function $f(x) = \mathrm{e}^x$ by noise of magnitude δ, and then computed its FE. As is evident, the discrete extension approximates f to an error of magnitude roughly δ, whereas for the continuous extension the error is of magnitude $\approx 10^6\delta$, as predicted by Table 1.

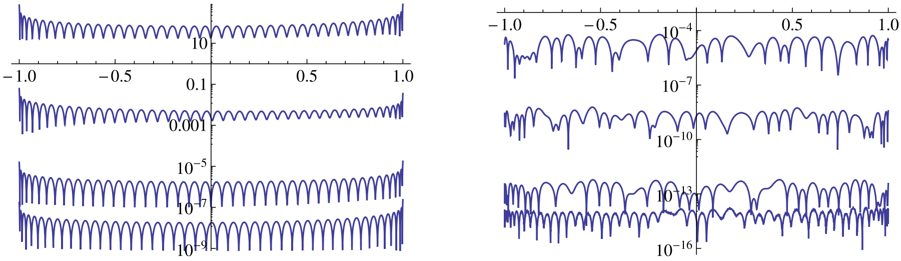


Fig. 8 The error $|f(x) - f_N(x)|$ against x, where $f_N = G_N(f)$ (*left*) or $f_N = \tilde{G}_N(f)$ (*right*), for $N = 30$, $T = 2$ and $f(x) = \mathrm{e}^x$, with noise at amplitudes $\delta = 10^{-4}, 10^{-8}, 10^{-12}, 0$

5 Fourier Extensions from Equispaced Data

We now turn our attention to the problem of computing FEs when only equispaced data is prescribed. As discussed in Sect. 1, a theorem of Platte, Trefethen and Kuijlaars [28] states that any exponentially convergent method for this problem must also be exponentially ill-conditioned (see Sect. 5.4 for the precise result). However, as we show in this section, FEs give rise to a method, the so-called *equispaced Fourier extension*, that allows this barrier to be circumvented to a substantial extent. Namely, it achieves rapid convergence in a numerically stable manner.

5.1 The Equispaced Fourier Extension

Let

$$x_n = \frac{n}{M}, \quad n = -M, \dots, M, \tag{5.1}$$

be a set of $2M + 1$ equispaced points in $[-1, 1]$, where $M \geq N$. We define the *equispaced Fourier extension* of a function $f \in \mathrm{L}^\infty[-1, 1]$ by

$$F_{N,M}(f) := \operatorname*{argmin}_{\phi \in \mathcal{G}_N} \sum_{|n| \leq M} \left| f(x_n) - \phi(x_n) \right|^2. \tag{5.2}$$

If $F_{N,M}(f) = \sum_{|n| \leq N} a_n \phi_n$, then the vector $a = (a_{-N}, \dots, a_N)^\top$ is the least squares solution to $\bar{A}a \approx \bar{b}$, where $\bar{A} \in \mathbb{C}^{(2M+1)\times(2N+1)}$ has (n, m)th entry $\frac{1}{\sqrt{M+1/2}} \phi_m(x_n)$ and $\bar{b}$ has nth entry $\frac{1}{\sqrt{M+1/2}} f(x_n)$.

Note that $F_{N,M}(f)$, as defined by (5.2), is (up to minor changes of parameters/notation) identical to the extensions considered in the previous papers [5, 8, 12, 24, 25] on equispaced FEs.

5.2 The Exact Equispaced Fourier Extension

Consider first the case $M = N$. Then $F_{N,N}(f)$ is equivalent to polynomial interpolation in z:

Proposition 5.1 *Let $F_{N,N}(f) = f_N = f_{\mathrm{e},N} + f_{\mathrm{o},N} \in \mathcal{G}_N$ be defined by* (5.2) *with $N = M$ and let $h_{i,N}(z)$ be given by* (2.4). *Then $h_{i,N}(z)$, $i = 1, 2$ is the $(N + 1 - i)$th degree polynomial interpolant of $h_i(z)$ at the nodes $\{z_n\}_{n=i-1}^{N} \subseteq [-1, 1]$, where*

$$z_n = m(x_n) = 2\frac{\cos(\frac{n\pi}{NT}) - c(T)}{1 - c(T)} - 1, \quad n = 0, \dots, N. \tag{5.3}$$

This proposition allows us to analyze the theoretical convergence/divergence of $F_{N,N}(f)$ using standard results on polynomial interpolation. Recall that associated with a set of nodes $\{z_n\}_{n=0}^{N}$ is a *node density function* $\mu(z)$, i.e. a function such that (i) $\int_{-1}^{1} \mu(z)\,\mathrm{d}z = 1$ and (ii) each small interval $[z, z + h]$ contains a total of $N\mu(z)h$ nodes for large N [18]. In the case of (5.3) we have

Lemma 5.2 *The nodes* (5.3) *have node density function $\mu(z) = T/(\pi\sqrt{(1 - z)(z - m(T))})$.*

Proof Note first that $\int_{-1}^{1} \mu(z)\,\mathrm{d}z = 1$. Now let $I = [z, z+h] \subseteq [-1, 1]$ be an interval. Then the node $z_n \in I$ if and only if $m^{-1}(z + h) \le x_n \le m^{-1}(z)$. Therefore, as $N \to \infty$, the proportion of nodes lying in I tends to $m^{-1}(z) - m^{-1}(z + h)$. Now suppose that $h \to 0$. Then

$$\begin{aligned} m^{-1}(z + h) &= \frac{T}{\pi} \arccos\left[c(T) + \frac{1 - c(T)}{2}(z + h + 1)\right] \\ &= m^{-1}(z) - \mu(z)h + \mathcal{O}\bigl(h^2\bigr). \end{aligned}$$

Thus $m^{-1}(z) - m^{-1}(z + h) = \mu(z)h + \mathcal{O}(h^2)$, as required. □

It is useful to consider the behavior of $\mu(z)$. When $z \to 1^-$, $\mu(z) \sim T/(\pi\sqrt{1 - z})$. On the other hand, μ is continuous at $z = -1$ with $\mu(-1) = \frac{T}{2\pi}\tan(\frac{\pi}{2T})$. Hence the nodes $\{z_n\}_{n=0}^{N}$ cluster quadratically near $z = 1$ and are linearly distributed near $z = -1$. It is well known that to avoid the Runge phenomenon in a polynomial interpolation scheme, it is essentially necessary for the nodes to cluster quadratically near both endpoints (as is the case with Chebyshev nodes) [18]. If this is not the case, one expects the Runge phenomenon: that is, divergence (at a geometric rate) of the interpolant for any function having a singularity in a certain complex region containing $[-1, 1]$ (the *Runge region* for the interpolation scheme). Since the nodes (5.3) do not exhibit the correct clustering at the endpoint $z = -1$, we consequently expect this behavior in the equispaced FE $F_{N,N}(f)$.

As it transpires, the corresponding Runge region $\mathcal{R} = \mathcal{R}(T)$ for $F_{N,N}$ can be defined in terms of the potential function $\phi(t) = -\int_{-1}^{1} \mu(z) \log|t - z|\,\mathrm{d}z + c$. Here c is an arbitrary constant. Standard polynomial interpolation theory [18] then gives

$$\mathcal{R}(T) = \bigl\{x \in \mathbb{C} : \phi\bigl(m(x)\bigr) = \phi(-1)\bigr\},$$

(observe that this is a subset of the complex x-plane). We note also that the convergence/divergence of $F_{N,N}(f)$ at a point x will be exponential at a rate

$e^{\phi(m(x_0))-\phi(m(x))}$, where x_0 is the limiting singularity of f. This follows from a general result on polynomial interpolation [18]. In particular, if f has a singularity in $\mathcal{R}(T)$, then there will be some points $x \in [-1,1]$ for which $F_{N,N}(f)$ diverges.

We next discuss two approaches for overcoming the Runge phenomenon in $F_{N,N}(f)$.

5.2.1 Overcoming the Runge Phenomenon I: Varying T

One way to attempt to overcome (or, at least, mitigate) the Runge phenomenon observed above is to vary the parameter T. Note that:

Lemma 5.3 *The Runge region $\mathcal{R}(T)$ satisfies $\mathcal{R}(T) \to [-1,1]$ as $T \to 1^+$, and $\mathcal{R}(T) \to \mathcal{R}$ as $T \to \infty$, where $\mathcal{R}$ is the Runge region for equispaced polynomial interpolation.*

Proof Suppose first that $T \to 1^+$. Since $m(T) \sim -1$, we have $\mu(z) \sim \frac{1}{\pi\sqrt{1-z^2}}$. The right-hand side is the potential function for Chebyshev interpolation, and thus the first result follows.

For the second result, we first recall that $\phi(m(x)) = -\int_{-1}^{1} \mu(z) \log|m(x)-z| \, \mathrm{d}z$. Define the change of variable $z = m(s)$. Since $m'(s) = -1/\mu(m(s))$ we have

$$\phi\bigl(m(x)\bigr) = -\int_0^1 \log\bigl|m(x)-m(s)\bigr| \, \mathrm{d}s.$$

Note that

$$\begin{aligned} m(x)-m(s) &= \frac{\cos\frac{\pi x}{T} - \cos\frac{\pi s}{T}}{\sin^2\frac{\pi}{2T}} \\ &= -\frac{2\sin\frac{\pi(x-s)}{2T}\sin\frac{\pi(x+s)}{2T}}{\sin^2\frac{\pi}{2T}} \sim -2(x-s)(x+s), \quad T \to \infty. \end{aligned}$$

Therefore

$$\phi\bigl(m(x)\bigr) \sim -\int_{-1}^{1} \log|x-s| \, \mathrm{d}s + C, \quad T \to \infty,$$

which is the potential function of equispaced polynomial interpolation, as required. □

This lemma comes as no surprise. As $T \to 1^+$ for fixed N, the system $\{e^{\mathrm{i}\frac{n\pi}{T}\cdot}\}_{|n|\le N}$ tends to the standard Fourier basis on $[-1,1]$. The problem of equispaced interpolation with trigonometric polynomials is well-conditioned and convergent. On the other hand, when $T \to \infty$, the subspaces $\mathcal{C}_N$ and $\mathcal{S}_N$ both resemble spaces of algebraic polynomials in x. Thus, in the large T limit, $F_{N,N}(f)$ is an algebraic polynomial interpolant of f at equispaced nodes.

Since the Runge region $\mathcal{R}(T)$ can be made arbitrarily small by letting $T \to 1^+$, one way to overcome the Runge phenomenon is to vary T in the way described in

Sect. 2.4 and set $T = T(N;\epsilon)$. One could also take $T \approx 1$ fixed. However, this will always lead to a nontrivial Runge region, and consequently divergence of $F_{N,N}$ for some nonempty class of analytic functions.

5.2.2 *Overcoming the Runge Phenomenon II: Oversampling*

An alternative means to overcome the Runge phenomenon in $F_{N,M}(f)$ is to allow $M \geq N$. Oversampling is known to defeat the Runge phenomenon in equispaced polynomial interpolation [8, 9, 28], and the same is true in this context (see [5, 12] for previous discussions on oversampling for equispaced FEs).

It is now useful to introduce some notation. For nodes $\{x_n\}_{|n|\leq M}$ given by (5.1), let $(\cdot,\cdot)_M$ be the discrete bilinear form $(g,h)_M = \frac{1}{M+\frac{1}{2}} \sum_{|n|\leq M} g(x_n)\overline{h(x_n)}$, and denote the corresponding discrete semi-norm by $\|\cdot\|_M$. Much as before, we define the condition number of $F_{N,M}$ by

$$\kappa(F_{N,M}) = \sup\bigl\{\bigl\|F_{N,M}(b)\bigr\| : b \in \mathbb{C}^{2M+1},\ \|b\| = 1\bigr\}. \tag{5.4}$$

We now have:

Theorem 5.4 *Let $F_{N,M}(f)$ be given by* (5.2), *and suppose that*

$$D(N,M) = \sup\bigl\{\|\phi\| : \phi \in \mathcal{G}_N,\ \|\phi\|_M = 1\bigr\}, \tag{5.5}$$

then

$$\bigl\|f - F_{N,M}(f)\bigr\| \leq \sqrt{2}\bigl(1 + D(N,M)\bigr) \inf_{\phi\in\mathcal{G}_N} \|f-\phi\|_\infty.$$

Moreover, the condition number $\kappa(F_{N,M}) = D(N,M)$.

Proof For the sake of brevity, we omit the first part of the proof (a very similar argument is given in [9] for the case of polynomial interpolation). For the second part, we first notice that

$$\kappa(F_{N,M}) = \sup\bigl\{\bigl\|F_{N,M}(f)\bigr\| : f \in \mathrm{L}^\infty(-1,1),\ \|f\|_M = 1\bigr\}.$$

Since $F_{N,M}(\phi) = \phi$ for $\phi \in \mathcal{G}_N$ we have $\kappa(F_{N,M}) \geq D(N,M)$. Conversely, since $F_{N,M}(f) \in \mathcal{G}_N$, and since $F_{N,M}$ is an orthogonal projection with respect to the bilinear form $(\cdot,\cdot)_M$, we have $\|F_{N,M}(f)\| \leq D(N,M)\|F_{N,M}(f)\|_M \leq D(N,M)\|f\|_M$. Hence $\kappa(F_{N,M}) \leq D(N,M)$, and we get the result. □

This theorem implies that $F_{N,M}(f)$ will converge, regardless of the analyticity of f, provided M is chosen such that $D(N,M)$ is bounded. Note that this is always possible: $D(N,M) \to 1$ as $M \to \infty$ for fixed N since $\|\cdot\|_M$ is a Riemann sum approximation to $\|\cdot\|$ and $\mathcal{G}_N$ is finite-dimensional. Up to small algebraic factors in M and N, the quantity $D(N,M)$ is equivalent to

$$\tilde{D}(N,M) = \sup\bigl\{\|p\|_\infty : p \in \mathbb{P}_N,\ \bigl|p(z_n)\bigr| \leq 1,\ n = 0,\dots,M\bigr\}. \tag{5.6}$$

Note the meaning of $\tilde{D}(N, M)$: it informs us how large a polynomial of degree N can be on $[-1, 1]$ if that polynomial is bounded at the M points z_n. Unfortunately, numerical evidence suggests that

$$\alpha^{\frac{N^2}{M}} \leq \tilde{D}(N, M) \leq \beta^{\frac{N^2}{M}}, \tag{5.7}$$

for constants $\beta \geq \alpha > 1$. Thus one requires $M = \mathcal{O}(N^2)$ nodes for boundedness of $D(N, M)$. This is clearly less than ideal: it means that we require many more samples of f to compute its N-term equispaced FE. In particular, the exact equispaced FE $F_{N,M}(f)$ of an analytic function f will converge only root-exponentially fast in the number M of equispaced grid values.

Had the nodes $\{z_n\}_{n=0}^{M}$ clustered quadratically near $z = \pm 1$, then $M = \mathcal{O}(N)$ would be sufficient to ensure boundedness of $\tilde{D}(N, M)$. Note that when $N = M$, $\tilde{D}(N, M)$ is precisely the Lebesgue constant of polynomial interpolation. On the other hand, if $\{z_n\}_{n=0}^{M}$ were equispaced nodes on $[-1, 1]$ then (5.7) would coincide with a well-known result of Coppersmith and Rivlin [15]. The intuition for a bound of the form (5.7) for the nodes (5.3) comes from the fact that these nodes are linearly distributed near $z = -1$. Thus, at least near $z = -1$ they behave like equispaced nodes.

We remark that it is straightforward to show that the scaling $M = \mathcal{O}(N^2)$ is sufficient for boundedness of $\tilde{D}(N, M)$. This is based on Markov's inequality for polynomials. Necessity of this condition would follow directly from the lower bound in (5.7), provided (5.7) were shown to hold. It may be possible to adapt the proof of [15] to establish this result.

Since the scaling $M = \mathcal{O}(N^2)$ is undesirable, one can ask what happens when $M = \gamma N$ for some fixed oversampling parameter $\gamma \geq 1$. Using potential theory arguments, one can show that $\tilde{D}(N, \gamma N)$ grows exponentially in N (with the constant of this growth becoming smaller as γ increases), as predicted by the conjectured bound (5.7). In other words,

$$N^{-1} \log D(N, \gamma N) \sim \log c(\gamma; T), \quad N \to \infty, \tag{5.8}$$

for some $c(\gamma; T) > 1$.[1] In view of this behavior, Theorem 5.4 guarantees convergence of the FE (5.2), provided $\rho \geq c(\gamma; T)$, where ρ is as in Theorem 2.11. In other words, f needs to be analytic in the region $\mathcal{D}(c(\gamma; T))$ (recall $\mathcal{D}$ from Theorem 2.11) to ensure convergence. Therefore, one expects a Runge phenomenon whenever f has a complex singularity lying in the corresponding Runge region $\mathcal{R}(\gamma; T) = \mathcal{D}(c(\gamma; T))$. Naturally, a larger value of γ leads to a smaller (but still nontrivial) Runge region. However, regardless of the choice of γ, there will always be analytic functions for which one expects divergence of $F_{N,\gamma N}(f)$ (see [9] for a related discussion in the case of equispaced polynomial interpolation). Moreover, the mapping $f \mapsto F_{N,\gamma N}$

[1] The constant of growth was obtained in private communication with A. Kuijlaars. A closed expression (up to several integrals involving the potential function ϕ for the nodes z_n) can be found for $c(\gamma; T)$. We omit the full argument as it is rather lengthy, but note that it is based on standard results in potential theory. A general reference is [29].

will always be exponentially ill-conditioned for any fixed γ, since the condition number is precisely $D(N, \gamma N)$ (Theorem 5.4).

Primarily for later use, we now note that it is also possible to study the condition number of the equispaced FE matrix $\bar{A}$ in a similar way. Straightforward arguments show that

$$1/\sigma_{\min}(\bar{A}) = B(N, M), \quad B(N, M) = \sup\bigl\{\|\phi\|_{[-T,T]} : \phi \in \mathcal{G}_N, \|\phi\|_M = 1\bigr\}. \tag{5.9}$$

Using the fact that $1/\sigma_{\min}(A) = \sup\{\|\phi\|_{[-T,T]} : \phi \in \mathcal{G}_N, \|\phi\| = 1\}$, where A is the matrix of the continuous FE, one can show that

$$1/\sigma_{\min}(A) \lesssim B(N, M) \le D(N, M)/\sigma_{\min}(A),$$

(here we use $\lesssim$ to mean up to possible algebraic factors in N). Theorem 3.1 now shows that $\bar{A}$ is always exponentially ill-conditioned in N, regardless of $M \ge N$.

Much like the case of $D(N, M)$ and $\tilde{D}(N, M)$, one can show that the quantity $B(N, M)$ is, up to algebraic factors, equivalent to

$$\tilde{B}(N, M) = \sup\bigl\{\|p\|_{\infty,[m(T),1]} : p \in \mathbb{P}_N, \bigl|p(z_n)\bigr| \le 1, n = 0, \dots, M\bigr\}. \tag{5.10}$$

Potential theory can be used once more to determine the exact behavior of $\tilde{B}(N, \gamma N)$. In particular,

$$N^{-1} \log B(N, \gamma N) \sim d(\gamma; T), \quad N \to \infty, \tag{5.11}$$

for some constant $d(\gamma; T) \ge c(\gamma; T) > 1$.

5.2.3 Numerical Examples

In the previous section we established (up to the conjecture (5.7)) 6., 7. and 8. of Sect. 1. The main conclusion is as follows. In order to obtain a convergent FE in exact arithmetic using equispaced data one either needs to oversample quadratically (and thereby reduce the convergence rate to only root-exponential), or scale the extension parameter T suitably with N or both. However, recall from Sect. 4 that a FE obtained from a finite precision computation may differ quite dramatically from the corresponding infinite-precision extension. Is it therefore possible that the unpleasant effects described in the previous section may not be witnessed in finite precision? The answer transpires to be yes, and consequently FEs can safely be used for equispaced data, even in situations where divergence is expected in exact arithmetic.

To illustrate, consider the function $f(x) = \frac{1}{1+100x^2}$. When $T = 2$, this function has a singularity lying in the Runge region $\mathcal{R}(1; 2)$. The predicted divergence of its exact (i.e. infinite-precision) equispaced FE is shown in Fig. 9. Note that double oversampling also gives divergence, whilst with quadruple oversampling the singularity of f no longer lies in $\mathcal{R}(\gamma; T)$. We therefore witness geometric convergence, albeit at a very slow rate. This behavior is typical. Given a function f it is always possible to select the oversampling parameter γ in such a way that $F_{N,\gamma N}(f)$ converges geometrically. However, such a γ depends on f in a nontrivial manner (i.e. the location of the nearest complex singularity of f) and therefore cannot in practice be determined

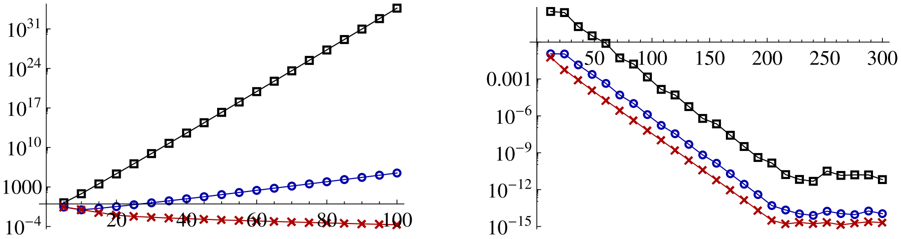


Fig. 9 The error $\|f - f_N\|_\infty$ against N for the equispaced FEs $f_N = F_{N,\gamma N}(f)$ (*left*) and $f_N = G_{N,\gamma N}(f)$ (*right*) of $f(x) = \frac{1}{1+100x^2}$ with oversampling factor $\gamma = 1, 2, 4$ (*squares*, *circles* and *crosses*) and $T = 2$

from the given data. Note that this phenomenon—namely, the fact that careful tuning of a particular parameter in a function-dependent way allows geometric convergence to be restored—is also seen in other methods for approximating functions to high accuracy, such as the Gegenbauer reconstruction technique [20, 21] (see Boyd [7] for a description of the phenomenon) and polynomial least squares [9].

Fortunately, and unlike for these other methods, the situation changes completely for Fourier extensions when we carry out computations in finite precision. This is shown in Fig. 9. For all choices of γ used, the finite precision FE, which we denote $G_{N,\gamma N}(f)$, converges geometrically fast, and there is no drift in the error once the best achievable accuracy is attained. Note that oversampling by a constant factor improves the approximation, but in all cases we still witness convergence. In particular, no careful selection of γ, such as that discussed above, appears to be necessary in finite precision.

5.3 The Numerical Equispaced Fourier Extension

We now explain these results by analyzing the numerical equispaced FE. Proceeding as in Sect. 4.2 we shall consider the truncated SVD approximation, which we denote $H_{N,M,\epsilon}(f)$. Note that a similar analysis has also recently been presented in [24]; see Remark 5.10 for further details.

Let $\Phi_n \in \mathcal{G}_N$ be the function corresponding to the right singular vector v_n of the matrix $\bar{A}$. Write $\mathcal{G}_{N,M,\epsilon} = \operatorname{span}\{\Phi_n : \sigma_n > \epsilon\}$ and $\mathcal{G}^{\perp}_{N,M,\epsilon} = \operatorname{span}\{\Phi_n : \sigma_n \leq \epsilon\}$, and note that $H_{N,M,\epsilon}$ is the orthogonal projection onto $\mathcal{G}_{N,M,\epsilon}$ with respect to $(\cdot,\cdot)_M$. Since $(\Phi_n, \Phi_m)_M = \sigma_n^2 \delta_{n,m}$, we have

$$H_{N,M,\epsilon}(f) = \sum_{n:\sigma_n > \epsilon} \frac{1}{\sigma_n^2}(f, \Phi_n)_M \Phi_n. \tag{5.12}$$

Our main result is as follows:

Theorem 5.5 *Let* $f \in \mathrm{L}^\infty(-1,1)$ *and* $H_{N,M,\epsilon}(f)$ *be given by* (5.12). *Then*

$$\begin{aligned}\big\|f - H_{N,M,\epsilon}(f)\big\| \leq{}& \sqrt{2}\big(1 + C_1(N,M;T,\epsilon)\big)\|f-\phi\|_\infty \\ &+ C_2(N,M;T,\epsilon)\|\phi\|_{[-T,T]}, \quad \forall \phi \in \mathcal{G}_N,\end{aligned} \tag{5.13}$$

and

$$\|a_\epsilon\| = \big\|H_{N,M,\epsilon}(f)\big\|_{[-T,T]} \le \frac{\sqrt{2}}{\epsilon}\|f-\phi\|_\infty + \|\phi\|_{[-T,T]}, \quad \forall \phi \in \mathcal{G}_N, \tag{5.14}$$

where

$$C_1(N,M;T,\epsilon) = \sup_{\substack{\phi\in\mathcal{G}_{N,M,\epsilon}\\ \phi\neq 0}} \left\{\frac{\|\phi\|}{\|\phi\|_M}\right\}, \qquad C_2(N,M;T,\epsilon) = \sup_{\substack{\phi\in\mathcal{G}^\perp_{N,M,\epsilon}\\ \phi\neq 0}} \left\{\frac{\|\phi\|}{\|\phi\|_{[-T,T]}}\right\}. \tag{5.15}$$

Proof Let $\phi \in \mathcal{G}_N$. Then

$$\big\|f - H_{N,M,\epsilon}(f)\big\| \le \|f-\phi\| + \big\|H_{N,M,\epsilon}(f-\phi)\big\| + \big\|\phi - H_{N,M,\epsilon}(\phi)\big\|. \tag{5.16}$$

Consider the second term. By definition of $C_1(N,M;T,\epsilon)$,

$$\big\|H_{N,M,\epsilon}(f-\phi)\big\| \le C_1(N,M,\epsilon)\big\|H_{N,M,\epsilon}(f-\phi)\big\|_M \le C_1(N,M,\epsilon)\|f-\phi\|_M,$$

where the second inequality follows from the fact that $H_{N,M,\epsilon}$ is an orthogonal projection with respect to $(\cdot,\cdot)_M$. Noting that $\|g\|, \|g\|_M \le \sqrt{2}\|g\|_\infty$ for any function $g \in \mathrm{L}^\infty(-1,1)$ now gives the corresponding term in (5.13). The bound for the third term of (5.16) follows immediately from the definition of $C_2(N,M;T,\epsilon)$ and the inequality $\|\phi - H_{N,M,\epsilon}(\phi)\|_{[-T,T]} \le \|\phi\|_{[-T,T]}$.

For (5.14), we first write $\|H_{N,M,\epsilon}(f)\|_{[-T,T]} \le \|H_{N,M,\epsilon}(f-\phi)\|_{[-T,T]} + \|H_{N,M,\epsilon}(\phi)\|_{[-T,T]}$. Observe that for any $g \in \mathrm{L}^\infty(-1,1)$ we have

$$\begin{aligned}\big\|H_{N,M,\epsilon}(g)\big\|^2_{[-T,T]} &= \sum_{n:\sigma_n>\epsilon} \frac{1}{\sigma_n^4}\big|(g,\Phi_n)_M\big|^2 \\ &\le \frac{1}{\epsilon^2}\big\|H_{N,M,\epsilon}(g)\big\|^2_M \le \frac{1}{\epsilon^2}\|g\|^2_M \le \frac{2}{\epsilon^2}\|g\|^2_\infty.\end{aligned}$$

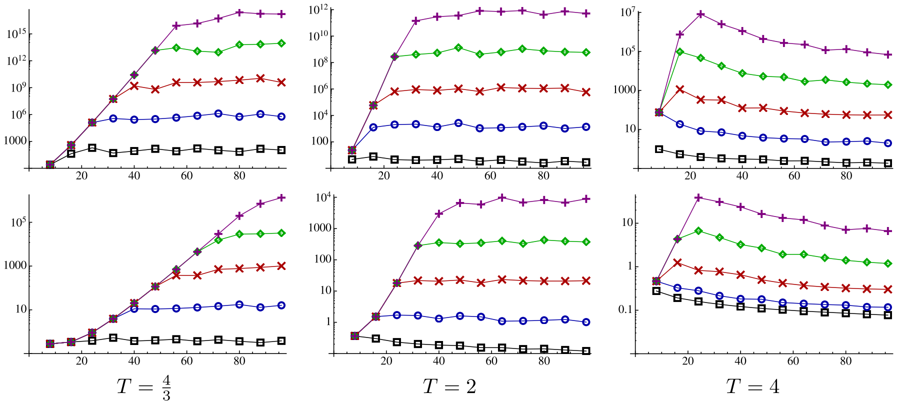
Also, $\|H_{N,M,\epsilon}(\phi)\|_{[-T,T]} \le \|\phi\|_{[-T,T]}$ for $\phi \in \mathcal{G}_N$. Setting $g = f - \phi$ and combining these two bounds now gives (5.14). □

Corollary 5.6 *If* $f \in \mathrm{L}^\infty(-1,1)$ *then* $\|H_{N,M,\epsilon}(f)\| \le \sqrt{2}/\epsilon\|f\|_\infty$, $\forall N \in \mathbb{N}$, $M \ge N$. *Moreover, if* $f \in \mathrm{H}^1(-1,1)$, $\mathbb{T} = [-T,T)$ *is the* T*-torus and* $c_1(T) > 0$ *is as in Lemma* 2.5, *then*

$$\begin{aligned}\limsup_{\substack{N,M\to\infty\\ M\ge N}} \big\|H_{N,M,\epsilon}(f)\big\| &\le \inf\big\{\|\tilde{f}\|_{[-T,T]} : \tilde{f} \in \mathrm{H}^1(\mathbb{T}),\ \tilde{f}|_{[-1,1]} = f\big\} \\ &\le c_1(T)\|f\|_{\mathrm{H}^1(-1,1)}.\end{aligned}$$

Proof By (5.14), we have

$$\big\|H_{N,M,\epsilon}(f)\big\| \le \big\|H_{N,M,\epsilon}(f)\big\|_{[-T,T]} \le \frac{\sqrt{2}}{\epsilon}\|f-\phi\|_\infty + \|\phi\|_{[-T,T]}, \quad \forall\phi \in \mathcal{G}_N. \tag{5.17}$$



$T = \frac{4}{3}$ $T = 2$ $T = 4$

Fig. 10 The quantity $C_1(N, \gamma N; T, \epsilon)$ against N for $\gamma = 1$ (*top row*) or $\gamma = 2$ (*bottom row*) and $\epsilon = 10^{-6}, 10^{-12}, 10^{-18}, 10^{-24}, 10^{-30}$ (*squares*, *circles*, *crosses*, *diamonds* and *dashes*, respectively)

Setting $\phi = 0$ gives the first result. For the second, we let ϕ be the N-term Fourier series of $\tilde{f}$ on $\mathbb{T}$, so that $\|f - \phi\|_\infty \to 0$ as $N \to \infty$. The final inequality follows from Lemma 2.5. □

This corollary shows that the equispaced FE cannot suffer from a Runge phenomenon in finite precision, since it is bounded in N and M. This should come as no surprise. Divergence of $H_{N,M,\epsilon}(f)$ would imply unboundedness of the coefficients a_ϵ, a behavior which is prohibited by truncating the singular values of $\bar{A}$ at level ϵ. Note that this corollary actually shows a much stronger result, namely that $H_{N,M,\epsilon}(f)$ is bounded on the extended domain $[-T, T]$, not just on $[-1, 1]$.

Although this corollary demonstrates lack of divergence of $H_{N,M,\epsilon}(f)$, it says littles about its convergence besides the observation that $\|H_{N,M,\epsilon}(f)\|$ is asymptotically bounded by $\|f\|_{\mathrm{H}^1(-1,1)}$. To study convergence we shall use (5.13). For this we first need to understand the constants $C_i(N, M; T, \epsilon)$.

5.3.1 Behavior of $C_i(N, M; T, \epsilon)$

Although Theorem 5.5 holds for arbitrary $M \geq N$, we now focus on the case of linear oversampling, i.e. $M = \gamma N$ for some $\gamma \geq 1$.

Let $N_2(\gamma, T, \epsilon)$ be the largest N such that all the singular values of $\bar{A}$ are at least ϵ in magnitude:

$$N_2(\gamma, T, \epsilon) = \max\{N : \sigma_{\min}(\bar{A}) > \epsilon\}.$$

For $N \leq N_2(\gamma, T, \epsilon)$ we have $\mathcal{G}_{N,\gamma N,\epsilon} = \mathcal{G}_N$ and therefore $C_1(N, \gamma N; T, \epsilon) = D(N, \gamma N)$, where $D(N, M)$ is given by (5.5). Thus we witness exponential divergence of $C_1(N, \gamma N; T, \epsilon)$ at rate $c(\gamma; T)$, where $c(\gamma; T)$ is as in (5.8). This is shown numerically in Fig. 10.

However, once $N > N_2(\gamma, T, \epsilon)$ the numerical results in Fig. 10 indicate a completely different behavior: namely, $C_1(N, \gamma N; T, \epsilon)$ appears to be bounded. Although we have no proof of this fact, these results strongly suggest the following

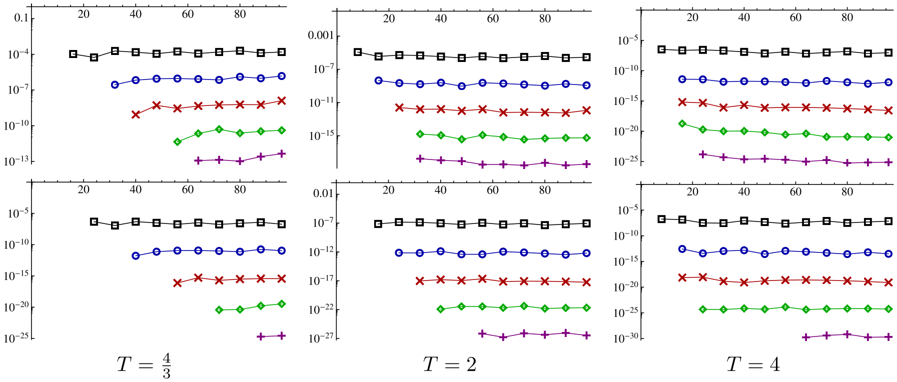


Fig. 11 The quantity $C_2(N,\gamma N;T,\epsilon)$ against N for $\gamma=1$ (*top row*) or $\gamma=2$ (*bottom row*) and $\epsilon=10^{-6},10^{-12},10^{-18},10^{-24},10^{-30}$ (*squares*, *circles*, *crosses*, *diamonds* and *dashes*, respectively)

conjecture:

$$C_1(N,\gamma N;T,\epsilon)\lesssim C_1(N_2,\gamma N_2;T,\epsilon)\sim c(\gamma;T)^{N_2},\quad \forall N\in\mathbb{N}. \tag{5.18}$$

In other words, $C_1(N,\gamma N;T,\epsilon)$ achieves its maximal value in N at $N\approx N_2$. Recalling (5.9) and (5.11), we note that

$$N_2(\gamma,T,\epsilon)\approx-\frac{\log\epsilon}{\log d(\gamma;T)}. \tag{5.19}$$

Thus, substituting this into bound (5.18) results in the following conjecture:

$$C_1(N,\gamma N;T,\epsilon)\lesssim\min\bigl\{c(\gamma;T)^N,\epsilon^{-\frac{\log c(\gamma;T)}{\log d(\gamma;T)}}\bigr\},\quad \forall N\in\mathbb{N}. \tag{5.20}$$

In particular, $C_1(N,\gamma N;T,\epsilon)$ is bounded for all N by some power of ϵ^{-1}. Importantly, this power cannot be too large. Note that $c(\gamma;T)\le d(\gamma;T)$, $\forall T>1$, since the maximum of a polynomial on $[m(T),1]$ is at least as large as its maximum on the smaller interval $[-1,1]$—compare (5.10) to (5.6). Therefore the ratio $\frac{\log c(\gamma;T)}{\log d(\gamma;T)}$ is at most one. Moreover, by varying either γ or T we may decrease this ratio to arbitrarily close to 1. We discuss this further in the next section.

The quantity $C_2(N,M;T,\epsilon)$ is harder to analyze, although clearly we have $C_2(N,M;T,\epsilon)=0$ when $N<N_2$. Figure 11 demonstrates that $C_2(N,\gamma N,\epsilon)$ is also bounded in N. Moreover, closer comparison with Fig. 10 suggests the existence of a bound of the form

$$C_2(N,\gamma N;T,\epsilon)\lesssim\epsilon C_1(N,\gamma N;T,\epsilon). \tag{5.21}$$

Once more, we have no proof of this observation.

Remark 5.7 The quantities $C_1(N,M;T,\epsilon)$ and $C_2(N,M;T,\epsilon)$ have the explicit expressions

$$C_1(N,M;T,\epsilon)=\sqrt{\bigl\|\bigl(S^\epsilon\bigr)^\dagger V^*AV\bigl(S^\epsilon\bigr)^\dagger\bigr\|},\qquad C_2(N,M;T,\epsilon)=\sqrt{\bigl\|\bigl(V^\epsilon\bigr)^*AV^\epsilon\bigr\|},$$

where A is the continuous FE matrix, USV^* is the singular value decomposition of the equispaced FE matrix $\bar{A}$, S^ϵ is formed by replacing the nth column of S by the zero vector whenever $\sigma_n \leq \epsilon$, and V^ϵ is formed by doing the same for columns of V corresponding to indices n with $\sigma_n > \epsilon$. These expressions were used to obtain the numerical results in Figs. 10 and 11. Computations were carried out with additional precision to avoid effects due to round-off.

5.3.2 Behavior of the Truncated SVD Fourier Extension

Combining the analysis of the previous section with Theorem 5.5, we now conjecture the bound

$$\left\| f - H_{N,\gamma N,\epsilon}(f) \right\| \leq C(\gamma, T, \epsilon)\big(\|f-\phi\|_\infty + \epsilon\|\phi\|_{[-T,T]}\big), \quad \forall \phi \in \mathcal{G}_N, \tag{5.22}$$

where $C(\gamma, T, \epsilon)$ is proportional to $\epsilon^{-a(\gamma;T)}$ and $a(\gamma;T)$ is given by

$$a(\gamma;T) = \frac{\log c(\gamma;T)}{\log d(\gamma;T)}. \tag{5.23}$$

This estimate allows us to understand the behavior of the numerical equispaced FE $G_{N,\gamma N}(f)$. When $N < N_2$ we have $G_{N,\gamma N}(f) = F_{N,\gamma N}(f)$ and therefore $G_{N,\gamma N}(f)$ will diverge geometrically fast in N whenever f has a singularity in the Runge region $\mathcal{R}(\gamma;T)$ (see Sect. 5.2.1). However, once N exceeds N_2, one obtains convergence. Indeed, setting $\phi = F_N(f)$ in (5.22), we find that the convergence is geometric up to the breakpoint N_1 (see (4.20)), and then, much as before, at least superalgebraic beyond that point. Note that the maximal achievable accuracy of order $C(\gamma, T, \epsilon)\epsilon \approx \epsilon^{1-a(\gamma;T)}$.

In summary, we have now identified the following convergence behavior for $H_{N,\gamma N,\epsilon}(f)$:

(i) $N < N_2(\gamma, T, \epsilon) \approx -\frac{\log\epsilon}{\log d(\gamma;T)}$. Geometric divergence/convergence of $H_{N,\gamma N,\epsilon}(f)$ at a rate of, at worst, $c(\gamma;T)/\rho$, where ρ is as in Theorem 2.11 and $c(\gamma;T)$ is given by (5.8).

(ii) $N_2(\gamma, T, \epsilon) \leq N < N_1(T, \epsilon) \approx -\frac{\log\epsilon}{\log E(T)}$. Geometric convergence at a rate of at least ρ.

(iii) $N = N_1(T, \epsilon)$. The error

$$\left\| f - H_{N,\gamma N,\epsilon}(f) \right\| \approx c_f \epsilon^{d_f - a(\gamma;T)},$$

where $a(\gamma;T)$ is as in (5.23) and $d_f = \frac{\log\rho}{\log E(T)} \in (0,1]$.

(iv) $N \geq N_1(\gamma, T)$. Superalgebraic convergence of $H_{N,\gamma N,\epsilon}(f)$ down to a maximal achievable accuracy proportional to $\epsilon^{1-a(\gamma;T)}$.

(This establishes 10. of Sect. 1.) Much as in the case of the discrete FE, we see that if f is analytic in $\mathcal{D}(E(T))$, and if c_f is not too large, then convergence stops at $N = N_1$ with maximal accuracy of order $c_f\epsilon^{1-a(\gamma;T)}$. Otherwise, we have a further regime of at least superalgebraic convergence before this accuracy is reached.

An important question is the role of the oversampling parameter γ in this convergence. We note:

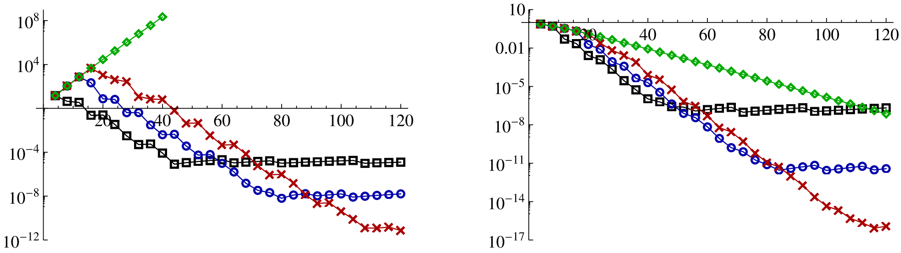


Fig. 12 Error against N for $H_{N,\gamma N,\epsilon}(f)$, where $f(x)=\frac{1}{1+16x^2}$, $T=2$, $\gamma=1$ (*left*) or $\gamma=2$ (*right*) and $\epsilon=10^{-6},10^{-12},10^{-18}$ (*squares*, *circles*, *crosses*). *Diamonds* correspond to the exact equispaced FE $F_{N,\gamma N}(f)$

Lemma 5.8 *Let $a(\gamma;T)$ be given by* (5.23). *Then $a(\gamma;T)$ satisfies $0\leq a(\gamma;T)\leq 1$ for all γ and T. Moreover, $a(\gamma;T)\to 0$ as $\gamma\to\infty$ for fixed T, and $a(\gamma;T)\to 0$ as $T\to\infty$ for fixed γ.*

Proof Note that $c(\gamma;T)\leq d(\gamma;T)$. Also $c(\gamma;T)\to 1$ and $d(\gamma;T)\to E(T)$ as $\gamma\to\infty$ for fixed T, and $d(\gamma;T)\to\infty$ as $T\to\infty$ for fixed γ, whereas $c(\gamma;T)$ is bounded. □

This lemma suggests that increasing γ will lead to a smaller constant $C(\gamma,T,\epsilon)$ in (5.22). In fact, numerical results (Figs. 10 and 11) indicate that using $T=2$ and $\gamma=2$ gives a bound of a little over 1 in magnitude for $\epsilon=10^{-14}$. Note that the effect of even just double oversampling is quite dramatic. Without oversampling (i.e. $\gamma=1$), the constant $C(\gamma,T,\epsilon)$ is approximately 10^4 in magnitude when $\epsilon=10^{-14}$ (see Figs. 10 and 11).

Let us make several further remarks. First, in practice the regime $N<N_1$ is typically very small—recall that N_1 is around 20 for $T=2$ (see Sect. 4.2.2)—and therefore one usually does not witness all three types of behavior in numerical examples. Second, as $\gamma\to\infty$, we have $N_2\to N_1$ (recall that $d(\gamma;T)\to E(T)$ as $\gamma\to\infty$). Thus, with a sufficient amount of oversampling, the regime (ii) will be arbitrarily small. On the other hand, oversampling decreases $c(\gamma;T)$, and therefore the rate of divergence in the regime (i) is also lessened by taking $\gamma>1$. Indeed, the numerical results in Fig. 12, as well as in Sect. 5.3.4 later, indicate that oversampling by a factor of 2 is typically sufficient in practice to mitigate the effects of divergence for most reasonable functions.

Figure 12 confirms these observations for the function $f(x)=\frac{1}{1+16x^2}$. For $\gamma=1$ the initial exponential divergence is quite noticeable. However, this effect largely vanishes when $\gamma=2$. Notice that a larger cutoff ϵ actually gives a smaller error initially, since there is a smaller regime of divergence. However, the maximal achievable accuracy is correspondingly lessened. We note also that maximal achievable accuracies for $\epsilon=10^{-6},10^{-12},10^{-18}$ are roughly 10^{-4}, 10^{-8} and 10^{-12}, respectively, when $\gamma=1$ and 10^{-7}, 10^{-12} and 10^{-16} when $\gamma=2$. These are in close agreement with the corresponding numerical values of $C_2(N,\gamma N;T,\epsilon)$ (see Fig. 11), as predicted by Theorem 5.5.

Table 2 The function $K(G_{N,\gamma N})$ against N with $T=2$ and $\gamma=1,2,4$

N	40	80	120	160	200
$\gamma=1$	2.37×10^4	3.50×10^4	2.24×10^4	2.47×10^4	1.93×10^4
$\gamma=2$	2.18×10^1	2.66×10^1	2.40×10^1	2.56×10^1	2.47×10^1
$\gamma=4$	8.03×10^0	1.05×10^1	1.23×10^1	1.39×10^1	1.54×10^1

Remark 5.9 A central conclusion of this section is that one requires a lower asymptotic scaling of M with N for the numerical equispaced FE than for its exact counterpart. Since $\mathcal{G}_{N,M,\epsilon}$ is a subset of $\mathcal{G}_N$, we clearly have $C_1(N,M;T,\epsilon)\le D(N,M)$, where $D(N,M)$ is given by (5.5). Hence quadratic scaling $M=\mathcal{O}(N^2)$ is sufficient (see Sect. 5.2.1) to ensure boundedness of $C_1(N,M;T,\epsilon)$, and one can make a similar argument for $C_2(N,M;T,\epsilon)$. However, Figs. 10 and 11 indicate that this condition is not necessary, and that one can get away with the much reduced scaling $M=\mathcal{O}(N)$ in practice.

This difference can be understood in terms of the singular values of $\bar{A}$. Recall that small singular values correspond to functions $\phi\in\mathcal{G}_N$ with $\|\phi\|_{[-T,T]}\gg\|\phi\|_M$. Now consider an arbitrary $\phi\in\mathcal{G}_N$. If the ratio $\|\phi\|/\|\phi\|_M$ is large, it suggests that ϕ lies approximately in the space $\mathcal{G}^{\perp}_{N,M,\epsilon}$ corresponding to small singular values. Hence, $\|\phi\|/\|\phi\|_M$ cannot be too large over $\phi\in\mathcal{G}_{N,M,\epsilon}$, and thus we see boundedness of $C_1(N,M,\epsilon)$, even when $D(N,M)$—the supremum of this ratio over the whole of $\mathcal{G}_N$—is unbounded.

Remark 5.10 A similar analysis of the equispaced FE, also based on truncated SVDs, was recently presented by M. Lyon in [24]. In particular, our expressions (5.13) and (5.22) are similar to equations (30) and (31) of [24]. Lyon also provides extensive numerical results for his analogues of the quantities $C_1(N,M;T,\epsilon)$ and $C_2(N,M;T,\epsilon)$, and describes a bound which is somewhat easier to use in computations. The main contributions of our analysis are the conjectured scaling of the constant $C(\gamma,T,\epsilon)$ in terms of ϵ, γ and T, the description and analysis of the breakpoints N_2 and N_1, and the differing convergence/divergence in the corresponding regions.

5.3.3 The Condition Number of the Numerical Equispaced FE

We now consider the condition number $\kappa(G_{N,M})$ (defined as in (5.4)) of the numerical equispaced extension. In Table 2 we plot $K(G_{N,\gamma N})$ against N, where $K(G_{N,M})$ is an upper bound for $\kappa(G_{N,M})$ defined analogously to (4.22). The results indicate numerical stability, and, as we expect, improved stability with more oversampling.

Besides oversampling it is also possible to improve stability by varying the extension parameter T. In Fig. 13 we give a contour plot of $K(G_{N,\gamma N})$ in the (γ,T)-plane. Evidently, increasing T improves stability. Recall, however, that a larger T corresponds to worse resolution power (see Sect. 2.4). Conversely, increasing γ also leads to worse resolution when measured in terms of the total number $M=\gamma N$ of

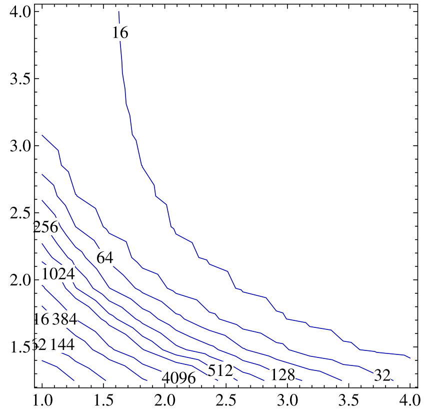


Fig. 13 Contour plot of the quantity $K(G_{N,\gamma N})$ against $1 \leq \gamma \leq 4$ and $1 < T \leq 4$ for $N = 200$

equispaced function values required. Hence a balance must be struck between the two quantities. Figure 13 suggests that $\gamma = T = 2$ is a reasonable choice in practice. Recall also that the choice $T = 2$ allows for fast computation of the equispaced FE (Remark 3.4), and hence is desirable to use in computations.

The behavior of the condition number can be investigated with the following theorem (the proof is similar to that of Theorem 4.7 and hence omitted):

Theorem 5.11 *The condition number $\kappa(H_{N,M,\epsilon})$ of the truncated SVD equispaced FE $H_{N,M,\epsilon}$ satisfies $\kappa(H_{N,M,\epsilon}) = C_1(N, M; T, \epsilon)$, where $C_1(N, M; T, \epsilon)$ is given by* (5.15).

From the analysis of Sect. 5.3.1 we conclude that $\kappa(H_{N,\gamma N,\epsilon}) \lesssim \epsilon^{-a(\gamma;T)}$, where $a(\gamma; T)$ is as in (5.23). Lemma 5.8 therefore shows that $\kappa(H_{N,\gamma N,\epsilon}) \lesssim 1$ as $\gamma \to \infty$ for fixed T, and $\kappa(H_{N,\gamma N,\epsilon}) \lesssim 1$ as $T \to \infty$ for fixed γ. This confirms the behavior described above.

5.3.4 Numerical Examples

In Fig. 14 we consider the equispaced FE for four test functions. In all cases we use $\gamma = 2$ and $T = 2$. As is evident, all choices of T give good, stable numerical results, with the best achievable accuracy being at least 10^{-12}. Robustness in the presence of noise is shown in Fig. 15. Observe that when $\gamma = 1$, noise of amplitude δ is magnified by around 10^5, in a manner consistent with Theorem 5.11. Conversely, with double oversampling, this factor drops to less than 10^2, again in agreement with Theorem 5.11.

5.4 Relation to the Theorem of Platte, Trefethen and Kuijlaars

We are now in a position to explain how FEs relate to the impossibility theorem of Platte, Trefethen and Kuijlaars [28]. A restatement of this theorem (with minor modifications to notation) is as follows:

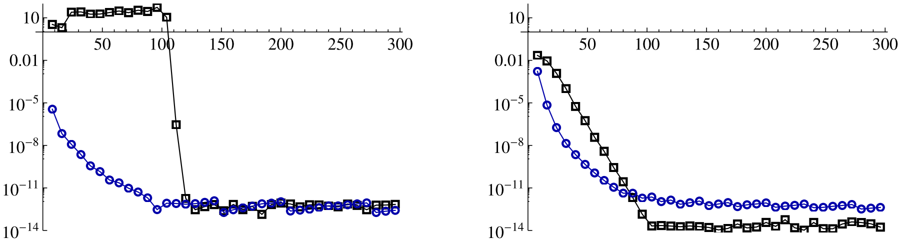


Fig. 14 The error $\|f - G_{N,\gamma N}(f)\|_\infty$, where $\gamma = 2$ and $T = 2$. *Left*: $f(x) = \mathrm{e}^{25\sqrt{5}\pi \mathrm{i} x}$ (*squares*), $f(x) = |x|^7$ (*circles*). *Right*: $f(x) = \frac{1}{1+25x^2}$ (*squares*), $f(x) = \frac{1}{8-7x}$ (*circles*)

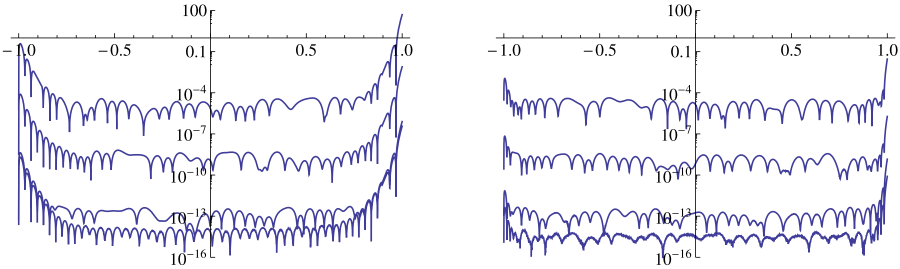


Fig. 15 The error $|f(x) - G_{N,\gamma N}(f)(x)|$ against x, where $\gamma = 1$ (*left*) or $\gamma = 2$ (*right*), for $N = 30$, $T = 2$ and $f(x) = \mathrm{e}^x$, with noise at amplitudes $\delta = 10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}, 0$

Theorem 5.12 [28] *Let F_M, $M \in \mathbb{N}$, be a sequence of approximations such that $F_M(f)$ depends only on the values of f on an equispaced grid of M points. Let $\mathcal{E} \subseteq \mathbb{C}$ be compact and suppose that there exist $C < \infty$, $\alpha > 1$ and $\tau \in (\frac{1}{2}, 1]$ such that*

$$\big\|f - F_M(f)\big\|_\infty \le C c_f \alpha^{-M^\tau}, \quad c_f = \max_{x\in\mathcal{E}} \big|f(x)\big|, \tag{5.24}$$

for all $M \in \mathbb{N}$ and all f that are continuous on $\mathcal{E}$ and analytic in its interior. Then there exists a $\beta > 1$ such that the condition numbers $\kappa(F_M) \ge \beta^{M^{2\tau-1}}$ for all sufficiently large M.

This theorem has two important consequences. First, any exponentially convergent method is also exponentially ill-conditioned. Second, the best possible convergence for a stable method is root-exponential in M. Note that the theorem is valid for all methods, both linear and nonlinear, that satisfy (5.24).

Consider now the exact equispaced Fourier extension $F_{N,M}$. As shown in Sect. 5.2, when $N = \mathcal{O}(\sqrt{M})$ this method is stable and root-exponentially convergent. Hence equispaced FEs in infinite precision *attain the maximal possible convergence rate for stable methods* satisfying the conditions of the theorem.

Now consider the numerical equispaced FE $G_{\eta M,M}$, where $0 < \eta \le 1$ is the reciprocal of the oversampling parameter γ used in the previous sections. We have shown that this approximation is stable, so at least one condition in Theorem 5.12 must be violated. Suppose that we take $\mathcal{E} = \mathcal{D}(E(T))$, for example. Then (5.22) shows that

$$\left\| f - G_{\eta M,M}(f) \right\|_\infty \lesssim c_f \epsilon^{-a(\eta^{-1};T)} \left(\left(E(T)^\eta \right)^{-M} + \epsilon \right). \tag{5.25}$$

The finite term ϵ in the brackets means that this approximation does not satisfy (5.24), and hence Theorem 5.12 does not apply. Recall that if c_f is small then (5.25) describes the full convergence behavior for all M. On the other hand, if c_f is large, or if $f \in \mathcal{D}(\rho)$ with $\rho < E(T)$, then the convergence is, after initial geometric convergence, at least superalgebraic down to the maximal achievable accuracy $\epsilon^{1-a(\eta^{-1};T)}$. This is also not in contradiction with the conditions of Theorem 5.12.

To summarize, equispaced FEs, when implemented in finite precision, possess both numerical stability and rapid convergence, and hence allow one to circumvent the impossibility theorem to an extent. In particular, for all functions $f \in \mathcal{D}(E(T))$ possessing small constants c_f, the approximations converge geometrically fast down to a maximal accuracy of order $\epsilon^{1-a(\eta^{-1};T)}$. In all other cases, the convergence is at least superalgebraic down to the same accuracy.

6 Conclusions and Challenges

We conclude by making the following remark. Extensive numerical experiments [5, 8, 12, 24, 25] have shown the effectiveness of FEs in approximating even badly behaved functions to high accuracy in a stable fashion. The purpose of this paper has been to provide analysis to explain these results. In particular, we have shown numerical stability for all three types of extensions considered, and analyzed their convergence. The reason for this robustness, despite the presence of exponentially ill-conditioned matrices, is due to the fact that the FE is a frame approximation and that for all functions f, even those with oscillations or large derivatives, there eventually exist coefficient vectors with small norms which approximate f to high accuracy.

The main outstanding theoretical challenge is to understand the constants $C_i(N, M; T, \epsilon)$ of the equispaced FE. In particular, we wish to show that linear scaling $M = \gamma N$ is sufficient to ensure boundedness of these constants in N, with a larger γ corresponding to a smaller bound. Note that the analysis of Sect. 5.2.1 implies the suboptimal result that $M = \mathcal{O}(N^2)$ is sufficient (Remark 5.9). It is also a relatively straightforward exercise to show that if $M = cN/\epsilon$ for suitable $c > 0$, then $C_i(N, M; T, \epsilon)$ is bounded. This is based on making rigorous the arguments given in Remark 5.9—we do not report it here for brevity's sake. Unfortunately, although this estimate gives the correct scaling $M = \mathcal{O}(N)$, it is wildly pessimistic. It implies that M should scale like $\approx 10^{16}N$, whereas the numerics in Sect. 5.3.1 indicate that $M = \gamma N$ is sufficient for *any* $\gamma \geq 1$.

One approach towards establishing a more satisfactory result is to perform a closer analysis of the singular values of the matrix $\bar{A}$. Some preliminary insight into this problem was given in [17]. Therein it was proved that (whenever $M = N$ and $2T \in \mathbb{N}$) the singular values cluster near zero and one, and the transition region is $\mathcal{O}(\log N)$ in width, much like for the prolate matrix A. Unfortunately, little is known outside of this result. There is no existing analysis for $\bar{A}$ akin to that of Slepian's for the prolate matrix—see [17] for a discussion. Note, however, that the normal form $B = \bar{A}^* \bar{A}$ has entries $B_{n,m} = \frac{\sin \frac{(n-m)\pi}{T}}{MT \sin \frac{(n-m)\pi}{MT}}$, and can therefore be viewed as a discretized version of the prolate matrix A. Indeed, $B \to A$ as $M \to \infty$ for fixed N. Given the similarities between the two matrices, there is potential for Slepian's analysis to be extended to this case. However, this remains an open problem.

Another issue is that of understanding how to choose the parameters T and γ in the case of the equispaced extension. As discussed in Sect. 2.4, the choice of T is reasonably clear for the continuous and discrete FEs (where there is no γ). If resolution of oscillatory functions is a concern, one should choose a small value of T (in particular, (2.18)). Otherwise, a good choice appears to be $T = 2$. However, for the equispaced FE, small T adversely affects stability (see Sect. 5.3.3). Hence it must be balanced by taking a larger value of the oversampling parameter γ, which has the effect of reducing the effective resolution power. In practice, however, a reasonable choice appears to be $T = \gamma = 2$. Investigating whether or not this is optimal is a topic for further investigation.

Acknowledgements The authors would like to thank John Boyd, Doug Cochran, Laurent Demanet, Anne Gelb, Anders Hansen, Arieh Iserles, Arno Kuijlaars, Mark Lyon, Nilima Nigam, Sheehan Olver, Rodrigo Platte, Jie Shen and Nick Trefethen for useful discussions and comments. They would also like to thank the anonymous referees for their constructive and helpful remarks.

References

1. B. Adcock, D. Huybrechs, On the resolution power of Fourier extensions for oscillatory functions. Technical Report TW597, Dept. Computer Science, K.U. Leuven, 2011.
2. N. Albin, O.P. Bruno, A spectral FC solver for the compressible Navier–Stokes equations in general domains. I: Explicit time-stepping, *J. Comput. Phys.* **230**(16), 6248–6270 (2011).
3. H. Bateman, *Higher Transcendental Functions*, vol. 2 (McGraw–Hill, New York, 1953).
4. J.P. Boyd, *Chebyshev and Fourier Spectral Methods* (Springer, Berlin, 1989).
5. J.P. Boyd, A comparison of numerical algorithms for Fourier extension of the first, second, and third kinds, *J. Comput. Phys.* **178**, 118–160 (2002).
6. J. Boyd, Fourier embedded domain methods: extending a function defined on an irregular region to a rectangle so that the extension is spatially periodic and C^∞, *Appl. Math. Comput.* **161**(2), 591–597 (2005).
7. J.P. Boyd, Trouble with Gegenbauer reconstruction for defeating Gibbs phenomenon: Runge phenomenon in the diagonal limit of Gegenbauer polynomial approximations, *J. Comput. Phys.* **204**(1), 253–264 (2005).
8. J.P. Boyd, J.R. Ong, Exponentially-convergent strategies for defeating the Runge phenomenon for the approximation of non-periodic functions. I. Single-interval schemes, *Commun. Comput. Phys.* **5**(2–4), 484–497 (2009).
9. J. Boyd, F. Xu, Divergence (Runge phenomenon) for least-squares polynomial approximation on an equispaced grid and Mock–Chebyshev subset interpolation, *Appl. Math. Comput.* **210**(1), 158–168 (2009).
10. O.P. Bruno, Fast, high-order, high-frequency integral methods for computational acoustics and electromagnetics, in *Topics in Computational Wave Propagation: Direct and Inverse Problems*, ed. by M. Ainsworth et al. Lecture Notes in Computational Science and Engineering, vol. 31 (Springer, Berlin, 2003), pp. 43–82.
11. O. Bruno, M. Lyon, High-order unconditionally stable FC–AD solvers for general smooth domains. I. Basic elements, *J. Comput. Phys.* **229**(6), 2009–2033 (2010).
12. O.P. Bruno, Y. Han, M.M. Pohlman, Accurate, high-order representation of complex three-dimensional surfaces via Fourier continuation analysis, *J. Comput. Phys.* **227**(2), 1094–1125 (2007).

13. C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang, *Spectral Methods: Fundamentals in Single Domains* (Springer, Berlin, 2006).
14. O. Christensen, *An Introduction to Frames and Riesz Bases* (Birkhauser, Basel, 2003).
15. D. Coppersmith, T. Rivlin, The growth of polynomials bounded at equally spaced points, *SIAM J. Math. Anal.* **23**, 970–983 (1992).
16. R.J. Duffin, A.C. Schaeffer, A class of nonharmonic Fourier series, *Trans. Am. Math. Soc.* **72**, 341–366 (1952).
17. A. Edelman, P. McCorquodale, S. Toledo, The future Fast Fourier Transform? *SIAM J. Sci. Comput.* **20**(3), 1094–1114 (1999).
18. B. Fornberg, *A Practical Guide to Pseudospectral Methods* (Cambridge University Press, Cambridge, 1996).
19. D. Gottlieb, S.A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, 1st edn. (SIAM, Philadelphia, 1977).
20. D. Gottlieb, C.-W. Shu, On the Gibbs' phenomenon and its resolution, *SIAM Rev.* **39**(4), 644–668 (1997).
21. D. Gottlieb, C.-W. Shu, A. Solomonoff, H. Vandeven, On the Gibbs phenomenon. I: Recovering exponential accuracy from the Fourier partial sum of a nonperiodic analytic function, *J. Comput. Appl. Math.* **43**(1–2), 91–98 (1992).
22. D. Huybrechs, On the Fourier extension of non-periodic functions, *SIAM J. Numer. Anal.* **47**(6), 4326–4355 (2010).
23. D. Kosloff, H. Tal-Ezer, A modified Chebyshev pseudospectral method with an $\mathcal{O}(N^{-1})$ time step restriction, *J. Comput. Phys.* **104**, 457–469 (1993).
24. M. Lyon, Approximation error in regularized SVD-based Fourier continuations, *Appl. Numer. Math.* **62**, 1790–1803 (2012).
25. M. Lyon, A fast algorithm for Fourier continuation, *SIAM J. Sci. Comput.* **33**(6), 3241–3260 (2012).
26. M. Lyon, O. Bruno, High-order unconditionally stable FC–AD solvers for general smooth domains. II. Elliptic, parabolic and hyperbolic PDEs; theoretical considerations, *J. Comput. Phys.* **229**(9), 3358–3381 (2010).
27. R. Pasquetti, M. Elghaoui, A spectral embedding method applied to the advection–diffusion equation, *J. Comput. Phys.* **125**, 464–476 (1996).
28. R. Platte, L.N. Trefethen, A. Kuijlaars, Impossibility of fast stable approximation of analytic functions from equispaced samples, *SIAM Rev.* **53**(2), 308–318 (2011).
29. T. Ransford, *Potential Theory in the Complex Plane* (Cambridge Univ. Press, Cambridge, 1995).
30. T.J. Rivlin, *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory* (Wiley, New York, 1990).
31. D. Slepian, Prolate spheroidal wave functions. Fourier analysis, and uncertainty. V: The discrete case, *Bell Syst. Tech. J.* **57**, 1371–1430 (1978).
32. L.N. Trefethen, D. Bau, *Numerical Linear Algebra* (SIAM, Philadelphia, 1997).
33. J. Varah, The prolate matrix, *Linear Algebra Appl.* **187**(1), 269–278 (1993).