From high oscillation to rapid approximation IV: Accelerating convergence

Daan Huybrechs
Department of Computer Science
Katholieke Universiteit Leuven
Celestijnenlaan 200A
B-3001 Leuven
Belgium

Arieh Iserles
DAMTP, Centre for Mathematical Sciences
University of Cambridge
Wilberforce Rd, Cambridge CB3 0WA
United Kingdom

Syvert P. Nørsett
Department of Mathematics
Norwegian University of Science and Technology
Trondheim N-7491
Norway

October 19, 2007

Abstract

Modified Fourier expansion is a powerful means for the approximation of non-periodic smooth functions in a univariate or multivariate setting. In the current paper we consider further enhancement of this approach by two techniques familiar from conventional Fourier analysis: polynomial subtraction and the hyperbolic cross. We demonstrate that, judiciously subtracting simpler functions dependent on linear combinations of derivatives along boundaries, it is possible to accelerate convergence a great deal and this procedure is considerably more efficient than in the case of conventional Fourier expansion. Moreover, examining the pattern of decay of coefficients in a multivariate setting, we demonstrate that most of them can be disregarded without any ill effect on quality of approximation.

1 Introduction

In this paper we continue work that commenced in (Iserles & Nørsett 2006*a*, Iserles & Nørsett 2006*b*, Iserles & Nørsett 2007), exploring approximation of functions in univariate and mul-

1

tivariate setting by modified Fourier expansions and its computational aspects.

Let $\Omega \subset \mathbb{R}^d$ be a bounded, closed, non-empty, simply-connected domain with piecewise-smooth boundary. (The assumption of simple connectedness is for convenience sake only, to render the use of the Stokes theorem straightforward, and can be easily lifted.) Modified Fourier expansion is concerned with expressing functions $f : \Omega \rightarrow \mathbb{R}$ as expansions in the eigenfunctions of the Laplace operator in Ω, equipped with Neumann boundary conditions (Iserles & Nørsett 2007).

Modified Fourier expansion features two important characteristics. Firstly, by virtue of the Neumann boundary conditions, it converges rapidly: for example (to which we will return), in the univariate setting it converges uniformly in any closed interval for C^3 functions and the speed of convergence using n expansion terms is $\mathcal{O}(n^{-2})$ within the interval and $\mathcal{O}(n^{-1})$ at the boundary (Iserles & Nørsett 2006a, Olver 2007). (In comparison, standard Fourier expansion for non-periodic functions converges like $\mathcal{O}(n^{-1})$ within the interval and fails to converge to the function values at the boundary.)

Second advantage of modified Fourier expansions is that the expansion coefficients are integrals of the function f, multiplied by rapidly oscillating Laplace–Neumann eigenfunctions. This brings them within realm of applicability of the very powerful methods for the numerical quadrature of highly oscillating integrals that have been developed in the last few years (Huybrechs & Vandewalle 2006, Iserles & Nørsett 2005, Olver 2006). Such methods allow for the approximate computation of the first n expansion coefficients in $\mathcal{O}(n)$ operations.

Modified Fourier expansions clearly are not a panacea for all computational problems. If Ω is a d-dimensional box and the function f is periodic, nothing can beat standard Fourier expansion and the Fast Fourier Transform. Even if f is not periodic, it is often preferable to use expansions in Chebyshev polynomials. Yet, we believe that the concept of modified Fourier expansions is powerful enough to confer real advantage in many situations and deserves further enquiry. In particular, we believe it worthwhile to explore ideas how to accelerate the convergence of modified Fourier series and how to implement them with considerably smaller set of coefficients (and in considerably smaller number of operations) without any injury to their precision. This is the theme of the present paper.

To set the stage for the new material presented in the sequel, we review very briefly aspects of univariate modified Fourier expansions from (Iserles & Nørsett 2006a). Assuming without loss of generality that $\Omega = [-1, 1]$, Laplace–Neumann eigenfunctions are simply

$$\{\cos \pi n x \ : \ n \in \mathbb{Z}_+\} \qquad \text{and} \qquad \{\sin \pi(n - \tfrac{1}{2})x \ : \ n \in \mathbb{N}\}$$

and the *modified Fourier expansion* is

$$f(x) = \tfrac{1}{2}\hat{f}_0^C + \sum_{n=1}^{\infty}[\hat{f}_n^C \cos \pi n x + \hat{f}_n^S \sin \pi(n - \tfrac{1}{2})x], \qquad (1.1)$$

where

$$\hat{f}_n^C = \int_{-1}^{1} f(t) \cos \pi n t \, \mathrm{d}t, \qquad \hat{f}_n^S = \int_{-1}^{1} f(t) \sin \pi(n - \tfrac{1}{2})t \, \mathrm{d}t.$$

Let $f \in \mathrm{C}^{\infty}[-1, 1]$. Using repeated integration by parts, it is easy to expand the *modified Fourier coefficients* $\hat{f}_n^C$ and $\hat{f}_n^S$ asymptotically for $n \gg 1$,

$$\hat{f}_n^C \sim \sum_{k=0}^{\infty} \frac{(-1)^{k+n}}{(\pi n)^{2k+2}}[f^{(2k+1)}(1) - f^{(2k+1)}(-1)], \qquad (1.2)$$

$$\hat{f}_n^S \sim \sum_{k=0}^{\infty} \frac{(-1)^{k+n+1}}{[\pi(n-\frac{1}{2})]^{2k+2}} [f^{(2k+1)}(1) + f^{(2k+1)}(-1)].\tag{1.3}$$

It follows at once that $\hat{f}_n^C, \hat{f}_n^S \sim \mathcal{O}(n^{-2})$ for $n \gg 1$. This is strikingly better than the conventional Fourier sine coefficients, that decay like $\mathcal{O}(n^{-1})$ and this is the main reason why (1.1) exhibits superior convergence properties (Iserles & Nørsett 2006a). However, arguably, the main importance of the asymptotic expansions (1.2) and (1.3) is numerical, rather than theoretical: they allow for very rapid calculation of quality approximations to the modified Fourier coefficients.

The simplest means to calculate $\hat{f}_n^C$ and $\hat{f}_n^S$ is based on the truncation of their respective asymptotic series. In that case we need to precompute the right number of odd derivatives at the endpoints ± 1. The procedure is remarkably effective even for fairly small values of n. Yet, it can be improved considerably, without requiring extra function evaluations, by the use of Filon-type formulæ (Iserles & Nørsett 2006a). This renders this approach competitive even for quite small values of n. The few values of n too small for such asymptotic-based techniques can be computed with little or no extra function evaluations using the 'exotic quadrature' approach explained in (Iserles & Nørsett 2006a).

Many ideas associated with the univariate modified Fourier expansion (1.1) scale up seamlessly to multivariate setting. In particular, an asymptotic expansion is available by employing repeatedly the Stokes theorem and can be used as a basis for the computation of expansion coefficients. The latter can be further enhanced by using Filon-type methods and 'exotic quadrature' (Iserles & Nørsett 2007). Of course, we take for granted that the full set of Laplace–Neumann eigenfunctions (of which rich theory exists: in particular, they form an orthogonal set closed in $L_2(\Omega)$) is explicitly known. This represents a tangible restriction on our approach.

One obvious generalisation of the univariate setting is to d-variate cubes using Cartesian products of the univariate eigenfunctions $\cos \pi n x$ and $\sin \pi (n - \frac{1}{2})x$. It is perhaps less straightforward than it seems and involves some fairly subtle constructs (Iserles & Nørsett 2007). Otherwise, even the list of simply-connected bivariate domains for which Laplace–Neumann eigenfunctions are comprehensively know is fairly restricted: it includes just ellipses and three types of triangles: equilateral, straight with two equal acute angles and straight with acute angles of $\frac{\pi}{6}$ and $\frac{\pi}{3}$. However, the situation is not as hopeless as this brief list might imply. Let Ω be any bivariate domain with piecewise-linear boundary (which need not be simply connected). To approximate $f \in C^{\infty}(\Omega)$ we tessellate Ω into triangles Δ_i, $i = 1, 2, \ldots, r$. Each Δ_i can be mapped affinely into, say, the unit equilateral triangle T. Denoting this mapping by φ_i, we approximate $f(\varphi_i^{-1})$ in T by truncated modified Fourier expansion g_i, say and map the outcome back to Δ_i. Thus, we are tiling Ω with approximations $g_i(\varphi_i)$, $i = 1, 2, \ldots, r$.

Detailed approximation by Laplace–Neumann eigenfunctions in an equilateral triangle will feature in a subsequent paper. At present it just suffices to point out that it is possible to represent the underlying eigenfunctions in a convenient form and derive handy formulæ and quadrature methods for expansion coefficients: everything scales up from the univariate case.

Unfortunately, so does the speed of convergence. While inside the triangle the n-term approximation converges like $\mathcal{O}(n^{-2})$, this drops down to $\mathcal{O}(n^{-1})$ on the boundary. This is particularly unwelcome when we employ the above tessellation of bivariate polygonal domains into triangles: the error is significantly larger across the (artificial) internal boundaries

than elsewhere. This is a manifestation of a wider problem implicit in our approach: although modified Fourier for non-periodic functions converges significantly faster than its conventional counterpart,[1] ideally we would have liked it to converge even faster, both inside the domain and on the boundary.

One obvious means to accelerate convergence is to replace Laplace–Neumann eigenfunctions by eigenfunctions of polyharmonic operators, subject to high-order Neumann boundary conditions. This is an approach presented for the univariate case in (Iserles & Nørsett 2006*b*). However, it is highly unlikely to be equally useful in a multivariate setting: even the simplest case, biharmonic eigenfunctions in a square, require non-elementary functions, and their practical implementation is currently an open problem.

An alternative is a technique already familiar in the context of standard Fourier expansions – indeed, one that might be construed as a competitor to modified Fourier: *polynomial subtraction* (Kantorovich & Krylov 1958, Roache 1978). The main idea is fairly simple. Thus, the nonperiodic function f is represented in the form $f = p + (f - p)$, where p is an $(2s - 1)$st-degree Hermite polynomial interpolation of f at ± 1,

$$p^{(i)}(\pm 1) = f^{(i)}(\pm 1), \qquad i = 0, \ldots, s - 1.$$

One then Fourier-expands the function $\tilde{f} = f - p$ and adds p. Since the first s derivatives of $\tilde{f}$ match at the endpoints, it is possible to prove that the (conventional) Fourier coefficients of $\tilde{f}$ decay like $\mathcal{O}(n^{-s-1})$. Moreover, the rate of convergence (for sufficiently smooth f) within the interval is $\mathcal{O}(n^{-s-1})$ and at the boundary it reduces to $\mathcal{O}(n^{-s})$.

As already mentioned in (Iserles & Nørsett 2006*a*), polynomial subtraction becomes substantially more effective when combined with modified Fourier expansions. The reason is clear from the asymptotic expansions (1.2) and (1.3): we need to subtract from f just the quadratic polynomial $p(x) = (\frac{1}{2}x - \frac{1}{4}x^2)f'(-1) + (\frac{1}{2}x + \frac{1}{4}x^2)f'(1)$ to force the modified Fourier coefficients of $\tilde{f}$ to decay like $\mathcal{O}(n^{-4})$. The same asymptotic decay within the realm of conventional Fourier expansions would have required matching f, f', f'' at the endpoints with quintic polynomial interpolation. The advantage of modified Fourier in this setting becomes even more pronounced once we require faster asymptotic decay.

Section 2 is devoted to detailed exploration of polynomial subtraction for modified Fourier expansions in an interval and in a d-dimensional box.

In Section 3 we explore another classical computational device which comes into its own in modified Fourier setting, the *hyperbolic cross* (Babenko 1960). We start from classical Fourier expansions in a unit square. In that case the (m, n) coefficient decays asymptotically like $\mathcal{O}(m^{-1}n^{-1})$. Therefore, once we truncate the series (as we are bound to in realistic computation), it is not necessary to retain *all* (m, n)th coefficients for $m, n \leq N$, say, where N is a sufficiently large integer, chosen to guarantee requisite precision. It is perfectly acceptable to require that $mn \leq M$ for some M (chosen according to similar criteria). This results in roughly $M \log_2 M$ terms in place of N^2 according to the naive approach. All this becomes even more advantageous with more variables.

Using hyperbolic cross within the context of modified Fourier expansions (ideally, in conjunction with polynomial subtraction) is considerably more natural. Of course, the asymptotic decay of the coefficients is more rapid, hence even fewer terms are needed. But the main difference is of a different kind. In the case of Fourier expansions integrals are inevitably computed with FFT, and the latter (in two variables) computes all the coefficients in a square: the

---

[1]Conventional Fourier approximations are, needless to say, defined only in boxes, certainly not in triangles.

large as well as the small. There is no easy way to excise the coefficients outside the hyperbolic cross! In our case, though, coefficients are computed with highly oscillatory quadrature, one-by-one. This means that we can easily restrict ourselves to the points in a lattice which lie inside a hyperbolic cross.

Of course, one clear advantage of multivariate modified Fourier expansions is that, unlike conventional Fourier series, they can be applied in more general domains, e.g. equilateral triangles. The economies implicit in the use of the hyperbolic cross are equally valid in this setting. We defer the discussion of polynomial subtraction and the hyperbolic cross in an equilateral triangle to another paper, where we will also consider other computational issues relevant in this setting.

2 Polynomial subtraction

2.1 The unit interval

Let $f \in C^\infty[-1, 1]$ be given: the theory can be extended in a transparent manner to $f \in C^s[-1, 1]$ for a suitable $s \geq 1$, with the minimal value of s changing according to setting, but the virtue of simplicity and clarity militates against generality for generality sake.

We seek a polynomial $p \in \mathbb{P}_{2r}$ such that $p(0) = 0$ and

$$p^{(2k+1)}(-1) = f^{(2k+1)}(-1), \quad p^{(2k+1)}(1) = f^{(2k+1)}(1), \qquad k = 0, 1, \ldots, r-1. \quad (2.1)$$

Our first observation is that we cannot take for granted the existence of such a polynomial. (If it exists it is clearly not unique, since we can add a constant to p without interfering with (2.1).) The conditions (2.1) are an example of *Birkhoff–Hermite polynomial interpolation* (Lorenz, Jetter & Riemenschneider 1983) and it is well known that a solution to such problems need not exist and must be carefully checked on a case-by-case basis.

Fortunately, insofar as (2.1) is concerned we can easily prove existence and, subject to a normalising condition, uniqueness.

Proposition 1 *For every $r \in \mathbb{N}$ there exists a unique polynomial p of degree $2r$ so that (2.1) holds and $p(0) = 0$.*

Proof   It is sufficient to show that there exists a unique $q \in \mathbb{P}_{2r-1}$ such that

$$q^{(2k)}(\pm 1) = f^{(2k+1)}(\pm 1), \qquad k = 0, \ldots, r-1$$

and set $p(x) = \int_0^x q(t)\mathrm{d}t$.

Let

$$q(x) = \sum_{l=0}^{2r-1} \frac{p_{l+1}}{l!} x^l.$$

Then the Birkhoff–Hermite interpolation conditions for q are equivalent to

$$\left.\begin{array}{rcl} \displaystyle\sum_{l=0}^{2r-2k-1} \frac{p_{l+2k+1}}{l!} & = & f^{(2k+1)}(1), \\ \displaystyle\sum_{l=0}^{2r-2k-1} (-1)^l \frac{p_{l+2k+1}}{l!} & = & f^{(2k+1)}(-1) \end{array}\right\} \qquad k = 0, 1, \ldots, r-1. \qquad (2.2)$$

Adding and subtracting the above equations, we convert them into two separate linear systems, each of r equations.

$$
\left.
\begin{array}{rcl}
\displaystyle\sum_{l=0}^{r-k-1} \frac{p_{2l+2k+1}}{(2l)!} &=& \frac{1}{2}[f^{(2k+1)}(1) + f^{(2k+1)}(-1)], \\[2ex]
\displaystyle\sum_{l=0}^{r-k-1} \frac{p_{2l+2k+2}}{(2l+1)!} &=& \frac{1}{2}[f^{(2k+1)}(1) - f^{(2k+1)}(-1)]
\end{array}
\right\}
\qquad k = 0, 1, \ldots, r-1.
$$

The two linear systems above (one for odd-indexed and the other for even-indexed coefficients) have nonsingular upper-triangular matrices. Hence they are nonsingular and possess unique solutions. This completes the proof. $\qquad\square$

Proposition 1 confirms the validity of polynomial subtraction in the setting of modified Fourier expansions in the unit interval $[-1, 1]$ (and, therefore, by linear translation, in any compact interval). To make the procedure explicit, we choose $r \in \mathbb{N}$, compute $f^{(2k+1)}(\pm 1)$ for $k = 0, \ldots, r-1$ and construct the interpolation polynomial p. In place of the expansion (1.1), we use

$$
f(x) = p(x) + \tfrac{1}{2}\hat{\tilde{f}}_0^C + \sum_{n=1}^{\infty} [\hat{\tilde{f}}_n^C \cos \pi n x + \hat{\tilde{f}}_n^S \sin \pi(n - \tfrac{1}{2})x], \tag{2.3}
$$

where

$$
\hat{\tilde{f}}_n^C = \int_{-1}^{1} [f(x) - p(x)] \cos \pi n x\, \mathrm{d}x, \qquad \hat{\tilde{f}}_n^S = \int_{-1}^{1} [f(x) - p(x)] \sin \pi(n - \tfrac{1}{2})x\, \mathrm{d}x.
$$

It follows at once from the asymptotic expansions (1.2) and (1.3) that

$$
\hat{\tilde{f}}_n^C \sim \sum_{k=r}^{\infty} \frac{(-1)^{k+n}}{(\pi n)^{2k+2}} [f^{(2k+1)}(1) - f^{(2k+1)}(-1)] = \mathcal{O}(n^{-2r-2}), \tag{2.4}
$$

$$
\hat{\tilde{f}}_n^S \sim \sum_{k=r}^{\infty} \frac{(-1)^{k+n+1}}{[\pi(n - \tfrac{1}{2})]^{2k+2}} [f^{(2k+1)}(1) - f^{(2k+1)}(-1)] = \mathcal{O}(n^{-2r-2}).
$$

This readily leads to the major observation of this subsection.

Theorem 2 *The asymptotic rate of decay of modified Fourier coefficients, once a polynomial $p \in \mathbb{P}_{2r}$ consistent with (2.1) is subtracted from f, is $\mathcal{O}(n^{-2r-2})$ for $n \gg 1$.*

The polynomial p can be derived solving the *normal equations* (2.2). An alternative, lending itself better toward detailed analysis and multivariate generalisation, is the use of *cardinal functions*. Thus, we seek polynomials $\theta_{n,\pm} \in \mathbb{P}_{2n}$ such that

$$
\theta_{n,\pm}^{(2l+1)}(\pm 1) = \begin{cases} 0, & l = 0, \ldots, n-2, \\ 1, & l = n-1, \end{cases} \qquad \theta_{n,\pm}^{(2l+1)}(\mp 1) = 0, \quad l = 0, \ldots, n-1.
$$

(Needless to say, $\theta_{n,\pm}^{(2l+1)} \equiv 0$ for $l \geq n$.) For example,

$$
\begin{aligned}
\theta_{1,\pm}(x) &= \tfrac{1}{2}(\pm\tfrac{1}{2}x^2 + x), \\
\theta_{2,\pm}(x) &= \tfrac{1}{24}(\pm\tfrac{1}{2}x^2 + x)(x^2 \pm 2x - 6), \\
\theta_{3,\pm}(x) &= \tfrac{1}{720}(\pm\tfrac{1}{2}x^2 + x)(x^4 \pm 4x^3 - 13x^2 \mp 34x + 75)
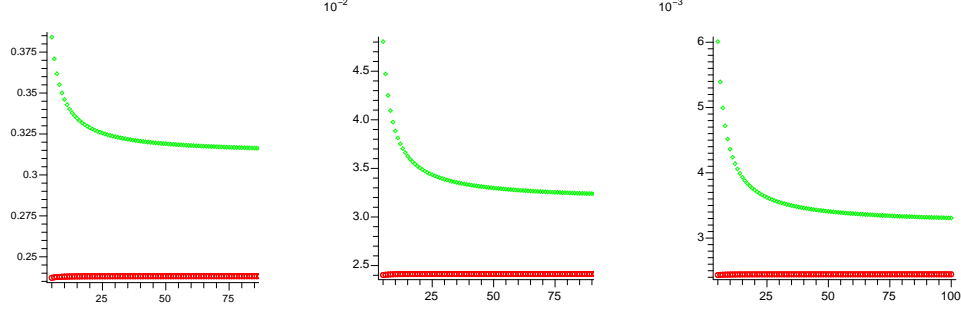\end{aligned}
$$

Figure 1: The magnitude of the coefficients (scaled by n^{2r+2}) for $r = 0$ (no polynomial subtraction), $r = 1$ and $r = 2$ for $f(x) = \mathrm{e}^x$.

and so on. In general, it is trivial to confirm that

$$\theta_{n,-}(x) = -\theta_{n,+}(-x), \qquad n \geq 1,$$

hence we need just the explicit form of $\theta_{n,+}$.

Proposition 3 *For every $n \in \mathbb{N}$ it is true that*

$$\theta_{n+1,+}(x) = a_n x + \tfrac{1}{2} b_n x^2 + \int_0^x (x - t)\theta_{n,+}(t)\mathrm{d}t, \qquad (2.5)$$

where

$$a_n = -\tfrac{1}{2} \int_0^1 [\theta_{n,+}(t) - \theta_{n,+}(-t)]\mathrm{d}t, \qquad b_n = -\tfrac{1}{2} \int_0^1 [\theta_{n,+}(t) + \theta_{n,+}(-t)]\mathrm{d}t.$$

Proof By induction. $\theta_{n,+} \in \mathbb{P}_{2n}$, therefore $\theta_{n+1,+} \in \mathbb{P}_{2n+2}$. Differentiation in (2.5), letting $x = \pm 1$ and substitution of the values of a_n and b_n confirm at once that

$$\theta'_{n+1,+}(\pm 1) = a_n \pm b_n + \int_0^{\pm 1} \theta_{n,+}(t)\mathrm{d}t = 0.$$

Moreover, it follows trivially from (2.5) that $\theta'''_{n+1,+} = \theta'_{n,+}$, hence, by induction on n, it is true that $\theta^{(2l+1)}_{n+1,+} = 0$, $l = 0, \ldots, n-1$, except that $\theta^{(2n+1)}_{n+1,+}(1) = \theta^{(2n-1)}_{n,+}(1) = 1$. Therefore $\theta_{n+1,+}$ obeys the interpolation conditions. □

In Fig. 1 we display the magnitude of the cosine (circles) and sine (diamonds) coefficients, scaled by n^{2r+2}, for $r = 0, 1, 2$ and the function $f(x) = \mathrm{e}^x$. As expected, the coefficients indeed decay like $\mathcal{O}(n^{-2r-2})$. What is perhaps surprising – yet is consistent with a great deal of further numerical experimentation – is how rapid is the onset of asymptotic behaviour.
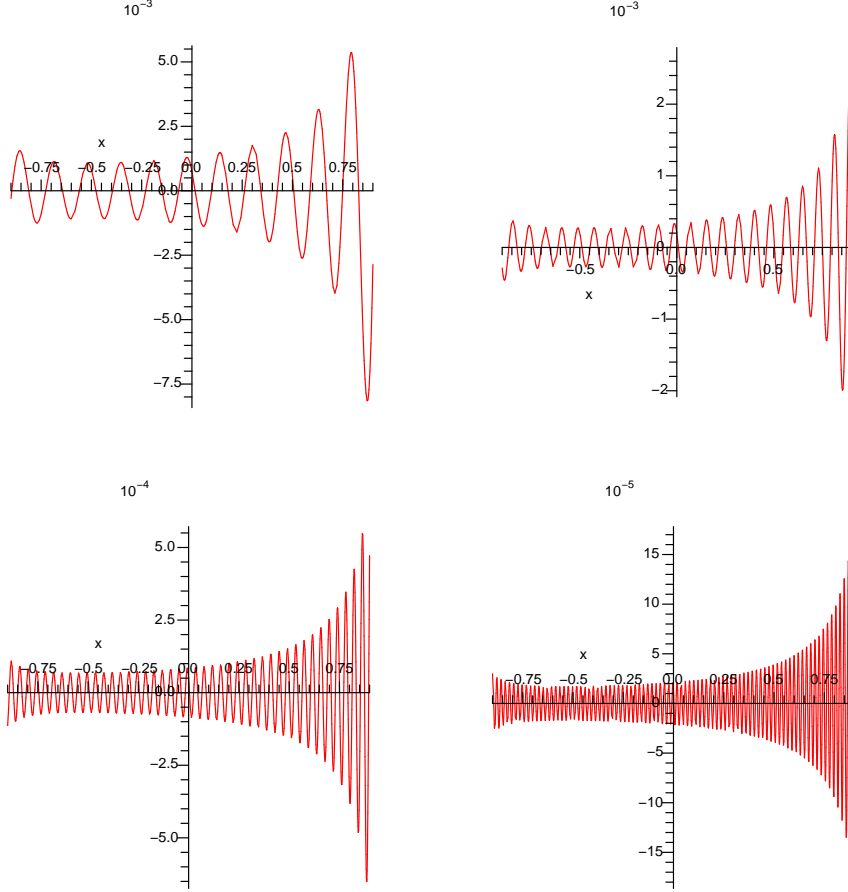
Figure 2: The error incurred in approximating $f(x) = e^x$ by modified Fourier expansions without polynomial subtraction, with 12, 24, 48 and 96 terms respectively.

This, of course, is of paramount importance when asymptotic formulæ are used as a basis for rapid numerical evaluation of expansion coefficients.

The theme of Figures 2–4 are pointwise errors incurred when e^x is approximated by modified Fourier expansions without polynomial subtraction (Fig. 2), with $r = 1$ (Fig. 3) and $r = 2$ (Fig. 4). The increase in precision is truly remarkable.

In the specific case $f(x) = e^x$ it is easy to derive coefficients explicitly for small r: letting $r = 0$ stand for the 'basic' expansion, without polynomial subtraction, we have

$$r = 0: \quad \hat{f}_0^C = (e - e^{-1}),$$

$$\hat{f}_m^C = \frac{(-1)^m (e - e^{-1})}{1 + \pi^2 m^2}, \quad \hat{f}_m^S = \frac{(-1)^{m+1}(e + e^{-1})}{1 + \pi^2 (m - \frac{1}{2})^2}, \quad m \in \mathbb{N},$$
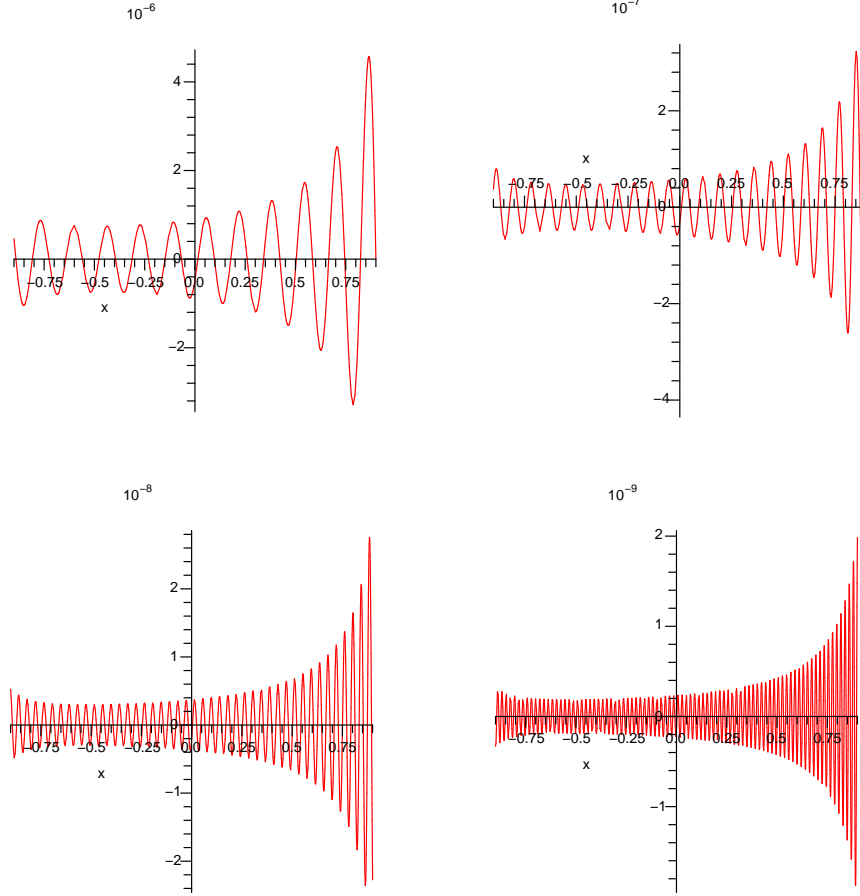
Figure 3: The error incurred in approximating $f(x) = e^x$ by modified Fourier expansions with polynomial subtraction, $r = 1$, with 12, 24, 48 and 96 terms respectively.

$$r = 1: \quad \hat{\tilde{f}}_0^C = \tfrac{5}{6}(e - e^{-1}),$$

$$\hat{\tilde{f}}_m^C = \frac{(-1)^{m+1}(e - e^{-1})}{\pi^2 m^2(1 + \pi^2 m^2)}, \quad \hat{\tilde{f}}_m^S = \frac{(-1)^m(e + e^{-1})}{\pi^2(m - \tfrac{1}{2})^2[1 + \pi^2(m - \tfrac{1}{2})^2]}, \quad m \in \mathbb{N},$$

$$r = 2: \quad \hat{\tilde{f}}_0^C = \tfrac{307}{360}(e - e^{-1}),$$

$$\hat{\tilde{f}}_m^C = \frac{(-1)^m(e - e^{-1})}{\pi^4 m^4(1 + \pi^2 m^2)}, \quad \hat{\tilde{f}}_m^S = \frac{(-1)^{m+1}(e + e^{-1})}{\pi^4(m - \tfrac{1}{2})^4[1 + \pi^2(m - \tfrac{1}{2})^2]}, \quad m \in \mathbb{N}.$$

This is fully consistent with the asymptotic rate of decay from Theorem 2. Moreover, it is
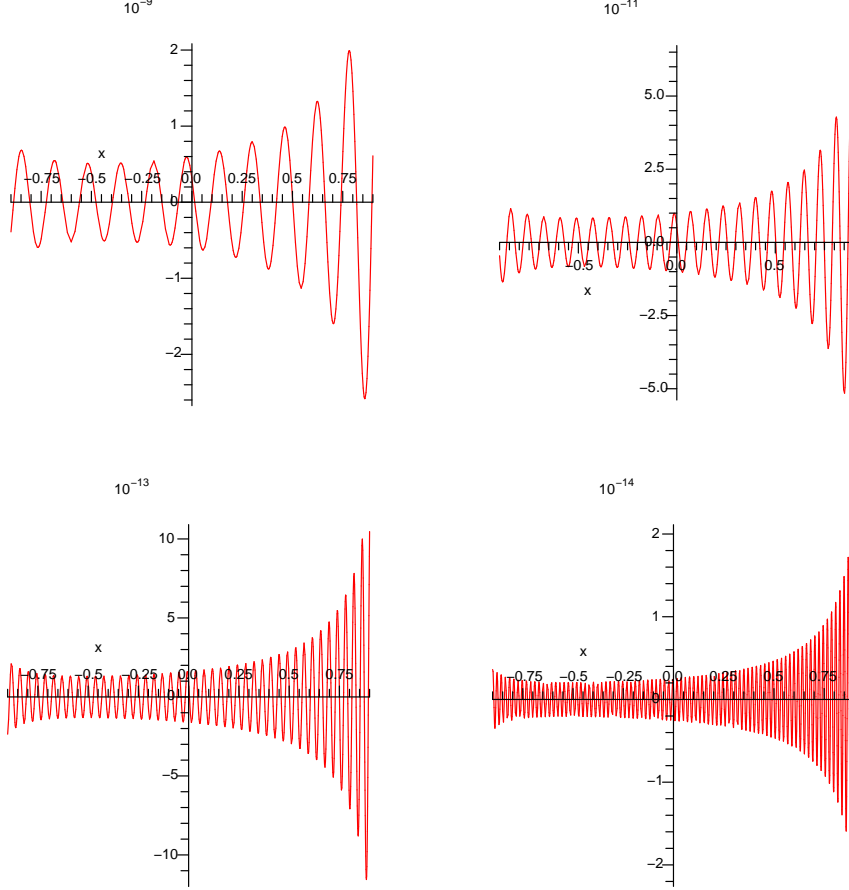
Figure 4: The error incurred in approximating $f(x) = e^x$ by modified Fourier expansions with polynomial subtraction, $r = 2$, with 12, 24, 48 and 96 terms respectively.

easy to confirm an obvious conjecture, namely that

$$\hat{\tilde{f}}_m^C = \frac{(-1)^{m+r}(e - e^{-1})}{(\pi m)^r[1 + (\pi m)^2]}, \quad \hat{\tilde{f}}_m^S = \frac{(-1)^{m+r+1}(e + e^{-1})}{[\pi(m - \frac{1}{2})]^r\{1 + [\pi(m - \frac{1}{2})]^2\}}, \quad m \in \mathbb{N},$$

for any $r \geq 0$ by summing up explicitly the asymptotic expansions (2.4).

The range of the above plots is the interval $\left[-\frac{9}{10}, \frac{9}{10}\right]$, thus we omit the vicinity of endpoints. The reason is that convergence at the endpoints is slower. In the standard modified Fourier case the error at ± 1 decays like $\mathcal{O}(n^{-1})$ (Iserles & Nørsett 2006a, Olver 2007) and it is easy to use the technique of Olver (2007) to prove that, for polynomial subtraction with $r \geq 1$ the error decay at the endpoints is $\mathcal{O}(n^{-2r-1})$ – a single unit less than inside the interval. This is fully confirmed by numerical results.

2.2 The square $[-1, 1]^2$

Detailed theory of modified Fourier expansions in d-dimensional boxes $[-1, 1]^d$ has been developed in (Iserles & Nørsett 2007). Since it involves a great deal of notation and terminology, designed to deal with many different types of coefficients, we concentrate in this subsection on the first nontrivial case, $d = 2$, observing that generalisation to more dimensions is straightforward in principle, delicate in practice.

A bivariate modified Fourier expansion of $f \in L_2([-1, 1]^2)$ is

$$\frac{1}{4}\hat{f}_{0,0}^{[0,0]} + \frac{1}{2}\sum_{n=1}^{\infty}[\hat{f}_{n,0}^{[0,0]}\cos(\pi nx) + \hat{f}_{0,n}^{[0,0]}\cos(\pi ny) + \hat{f}_{n,0}^{[1,0]}\sin(\pi(n - \tfrac{1}{2})x)$$

$$+ \hat{f}_{0,n}^{[0,1]}\sin(\pi(n - \tfrac{1}{2})y)] + \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty}[\hat{f}_{n_1,n_2}^{[0,0]}\cos(\pi n_1 x)\cos(\pi n_2 y) \qquad (2.6)$$

$$+ \hat{f}_{n_1,n_2}^{[1,0]}\sin(\pi(n_1 - \tfrac{1}{2})x)\cos(\pi n_2 y) + \hat{f}_{n_1,n_2}^{[0,1]}\cos(\pi n_1 x)\sin(\pi(n_2 - \tfrac{1}{2})y)$$

$$+ \hat{f}_{n_1,n_2}^{[1,1]}\sin(\pi(n_1 - \tfrac{1}{2})x)\sin(\pi(n_2 - \tfrac{1}{2})y)],$$

where

$$\hat{f}_{n_1,n_2}^{[0,0]} = \int_{-1}^{1}\int_{-1}^{1} f(x, y)\cos(\pi n_1 x)\cos(\pi n_2 y)\mathrm{d}x\mathrm{d}y,$$

$$\hat{f}_{n_1,n_2}^{[1,0]} = \int_{-1}^{1}\int_{-1}^{1} f(x, y)\sin(\pi(n_1 - \tfrac{1}{2})x)\cos(\pi n_2 y)\mathrm{d}x\mathrm{d}y,$$

$$\hat{f}_{n_1,n_2}^{[0,1]} = \int_{-1}^{1}\int_{-1}^{1} f(x, y)\cos(\pi n_1 x)\sin(\pi(n_2 - \tfrac{1}{2})y)\mathrm{d}x\mathrm{d}y,$$

$$\hat{f}_{n_1,n_2}^{[1,1]} = \int_{-1}^{1}\int_{-1}^{1} f(x, y)\sin(\pi(n_1 - \tfrac{1}{2})x)\sin(\pi(n_2 - \tfrac{1}{2})y)\mathrm{d}x\mathrm{d}y.$$

The modified Fourier coefficients above can be expanded asymptotically. A departure from the univariate case is that the coefficients now depend on two parameters, n_1 and n_2, both of which can grow large independently of each other. Hence, there is a two-dimensional quadrant of coefficients and, ideally, we want to accelerate the decay of modified Fourier coefficients in any direction where either n_1 or n_2 grows large – or both.

We commence the discussion with the easiest case, where n_1 and n_2 are both large. The first term in this expansion is $\mathcal{O}(n_1^{-2}n_2^{-2})$, the second is a linear combination of $\mathcal{O}(n_1^{-4}n_2^{-2})$ and $\mathcal{O}(n_1^{-2}n_2^{-4})$ terms and so on: we refer the reader to (Iserles & Nørsett 2007) for details. A natural generalisation of the framework from the last subsection to a bivariate setting is to subtract from f a multivariate polynomial p so that

$$f_{xy}(\pm 1, \pm 1) = p_{xy}(\pm 1, \pm 1). \qquad (2.7)$$

This means that we knock out the leading asymptotic term and, for large n_1, n_2, modified Fourier coefficients of $f - p$ decay like $\mathcal{O}(n_1^{-4}n_2^{-2}, n_1^{-2}n_2^{-4})$. Observing that the next terms in the expansion always depend on the partial derivatives of odd order in x and y, evaluated precisely at the vertices of the square, we can iterate this approach. Thus, in order to accelerate the decay of the coefficients further, we require in addition to (2.7)

$$f_{xxxy}(\pm 1, \pm 1) = p_{xxxy}(\pm 1, \pm 1) \quad \text{and} \quad f_{xyyy}(\pm 1, \pm 1) = p_{xyyy}(\pm 1, \pm 1),$$

and so on.

An important subset of bivariate modified Fourier coefficients are the 'boundary coefficients', when $\min\{n_1, n_2\} = 0$. These coefficients are non-oscillatory in one dimension and hence they exhibit different asymptotic expansion. Replacing $f(x)$ with $F(x) = \int_{-1}^{1} \tilde{f}(x, y)\mathrm{d}y$ in (1.2), it follows that

$$\hat{\tilde{f}}_{n,0}^{[0,0]} \sim \sum_{k=0}^{\infty} \frac{(-1)^{n+k}}{(\pi n)^{2k+2}} [F^{(2k+1)}(1) - F^{(2k+1)}(-1)].$$

Thus, to acquire rapid asymptotic decay for $n \gg 1$ we need $F'(1) = F'(-1)$ and we can easily prove that, once other 'boundary coefficients' are taken into account, we need

$$\int_{-1}^{1} f_x(\pm 1, y)\mathrm{d}y = \int_{-1}^{1} p_x(\pm 1, y)\mathrm{d}y, \qquad \int_{-1}^{1} f_y(x, \pm 1)\mathrm{d}x = \int_{-1}^{1} p_y(x, \pm 1)\mathrm{d}x, \quad (2.8)$$

extra four conditions on the polynomial p.

Imposing just conditions (2.7) and (2.8) on p does not yet achieve accelerated decay in all directions, as it is still true that $\hat{\tilde{f}}_{n_1,n_2}^{[0,0]} = \mathcal{O}(n_1^{-2})$ for large n_1, but small $n_2 > 0$. We can repeat the argument that led to (2.8) and replace $f(x)$ by

$$F(x) = \int_{-1}^{1} \tilde{f}(x, y) \cos(n_2 \pi y)\mathrm{d}y,$$

where n_2 may be small. It follows that we again need $F'(1) = F'(-1)$. More generally, to obtain rapid decay of all modified Fourier coefficients with large n_1 and fixed $n_2 \in \mathbb{N}$, this argument leads to the conditions

$$\int_{-1}^{1} f_x(\pm 1, y) \cos(\pi n_2 x)\mathrm{d}y = \int_{-1}^{1} p_x(\pm 1, y) \cos(\pi n_2 x)\mathrm{d}y,$$

$$\int_{-1}^{1} f_x(\pm 1, y) \sin(\pi(n_2 - \tfrac{1}{2})x)\mathrm{d}y = \int_{-1}^{1} p_x(\pm 1, y) \sin(\pi(n_2 - \tfrac{1}{2})x)\mathrm{d}y.$$

Thus, we find that p_x should reproduce precisely the *univariate* modified Fourier coefficients of f_x along the vertical edges of the boundary. A similar result follows for the case of large n_2 and fixed n_1, involving the two horizontal edges of $[-1, 1]^2$. Alternatively, all required conditions are automatically satisfied once we interpolate the normal derivative of f along the entire boundary,

$$\frac{\partial f}{\partial n}(x, y) = \frac{\partial p}{\partial n}(x, y), \qquad (x, y) \in \partial[-1, 1]^2. \qquad (2.9)$$

Any function p satisfying this condition also satisfies all conditions mentioned above. Note in particular that any sufficiently smooth function p satisfying (2.9) should also satisfy (2.7).

Condition (2.9) does not come as a surprise. A general expansion for modified Fourier coefficients, that converges in an asymptotic sense, was developed in (Iserles & Nørsett 2007),

$$\langle f, u \rangle \sim -\sum_{k=0}^{\infty} \frac{1}{(-\lambda)^{k+1}} \int_{\partial \Omega} \frac{\partial \Delta^k f(\boldsymbol{x})}{\partial n} u(\boldsymbol{x})\mathrm{d}S, \qquad \lambda \gg 1, \qquad (2.10)$$

based only on a repeated application of the Stokes theorem. Here, $u(\boldsymbol{x})$ is a Laplace–Neumann eigenfunction in Ω with the corresponding eigenvalue λ. In the case of a square, the eigenvalue behaves as $\lambda \sim n_1^2 + n_2^2$, and thus $\lambda \gg 1$ if $\max\{n_1, n_2\} \gg 1$. Matching normal derivatives along the boundary means the first term in (2.10) is cancelled. This is fully consistent with the objective of accelerated convergence whenever n_1, n_2 or both are large. At the same time, this expansion signposts the path of generalization towards both cancelling higher order asymptotic terms and extending acceleration to other bivariate sets Ω.

The interpolation of the normal derivative along the entire boundary might seem a tall order – even more so if we wish to speed asymptotic decay further by subtracting higher derivatives. Fortunately, this can be done fairly easily by employing techniques from computer-aided geometric design similar to the *Coons patches* (Farin 1997). Strictly speaking, the outcome is not *polynomial* subtraction, since p is no longer a polynomial, but this should cause no problems whatsoever in practice.

Recall the quadratic $\theta_1(x) = \theta_{1,+}(x)$ from the last subsection and note that $\theta_{1,-}(x) = -\theta_1(-x)$. We let

$$
\begin{aligned}
p(x,y) = &-\theta_1(-x)f_x(-1,y) + \theta_1(x)f_x(1,y) - \theta_1(-y)f_y(x,-1) + \theta_1(y)f_y(x,1) \\
&- [\theta_1(-x)\theta_1(-y)f_{xy}(-1,-1) - \theta_1(-x)\theta_1(y)f_{xy}(-1,1) \\
&- \theta_1(x)\theta_1(-y)f_{xy}(1,-1) + \theta_1(x)\theta_1(y)f_{xy}(1,1)].
\end{aligned}
\tag{2.11}
$$

Note that p 'fills in' values in $(-1,1)^2$, using solely information on f along the boundary. It is now easy to verify that

$$
p_x(\pm 1, y) = f_x(\pm 1, y), \qquad p_y(x, \pm 1) = f_y(x, \pm 1),
$$

therefore (2.9) is satisfied and we can use p to 'subtract' the influence of the leading asymptotic term in *all* types of coefficients.

There is an apparent element of cheating in using the above function p to accelerate convergence. After all, the entire point of approximation theory is to express (or approximate) complicated functions in terms of simpler functions: polynomials, trigonometric functions, splines etc. This is entirely consistent with polynomial subtraction, since then f is expressed as a polynomial plus trigonometric series, but not when p is allowed to be a linear combination of derivatives of f along the boundary. However, this is not a real problem. The functions $f_x(\pm 1, y)$ and $f_y(x, \pm 1)$ are univariate and we can use the theory of Subsection 2.1 to approximate them rapidly with 'proper' polynomial subtraction. In other words, our procedure (which generalises to higher dimensions) is to subtract from f a linear combination of lower-dimensional functions which are, in turn, approximated in a lower-dimensional box.

With greater generality, and bearing in mind expansion (2.10), in order to eliminate all $\mathcal{O}\left(n_1^{-2j_1}n_2^{-2j_2}\right)$ terms in the square with $j_1, j_2 \leq r$ we need to fit all normal derivatives of the form $\partial \Delta^j f / \partial n$ for $j = 0, \ldots, r-1$ along the boundary of the square. To this end we choose $r \in \mathbb{N}$ and let

$$
\begin{aligned}
p_r(x,y) = \sum_{k=1}^{r}[&-\theta_k(-x)\mathrm{D}_x^{2k-1}f(-1,y) + \theta_k(x)\mathrm{D}_x^{2k-1}f(1,y) - \theta_k(-y)\mathrm{D}_y^{2k-1}f(x,-1) \\
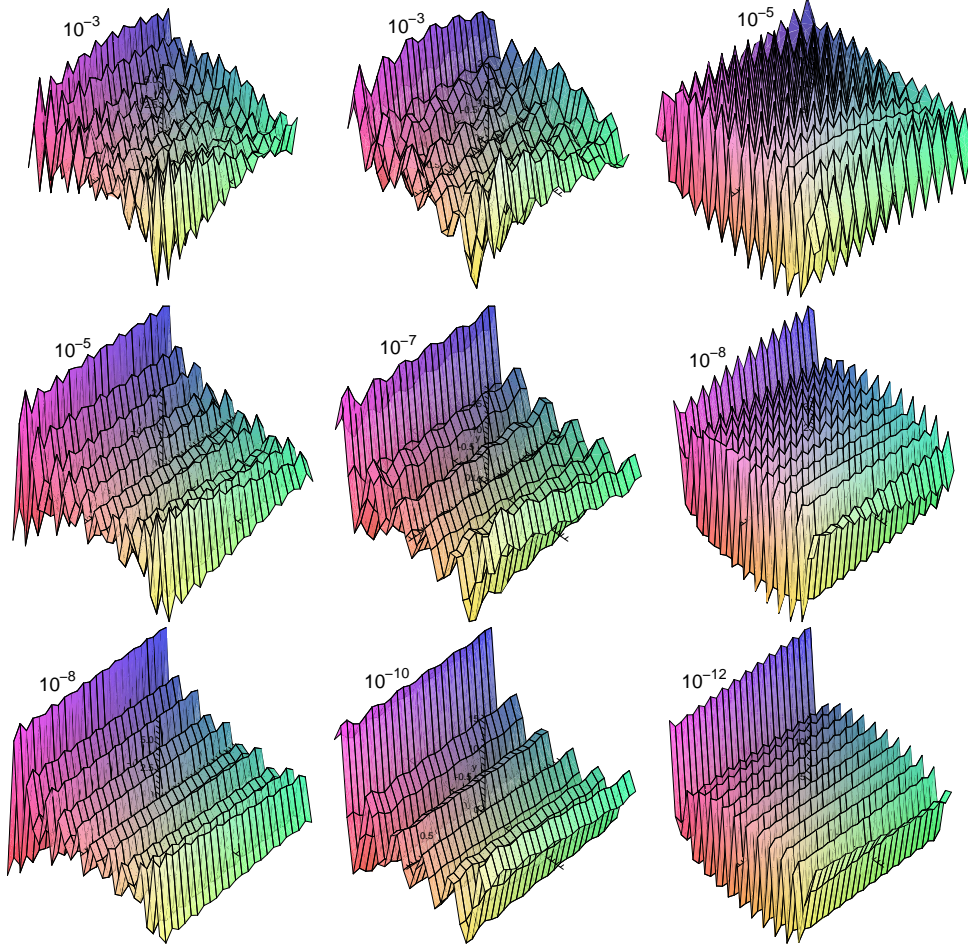&+ \theta_k(y)\mathrm{D}_y^{2k-1}f(x,1)]
\end{aligned}
$$

Figure 5: The error incurred in approximating $f(x, y) = \sin(1 + x - 2y)$ by modified Fourier expansions with $r = 0$ (top row), $r = 1$ and $r = 2$ (bottom row), truncating for $n_1, n_2 \leq 10$ (left column), $n_1, n_2 \leq 20$ (central column) and $n_1, n_2 \leq 40$ (right column).

$$-\sum_{k=1}^{r}\sum_{l=1}^{r}[\theta_k(-x)\theta_l(-y)\mathrm{D}_x^{2k-1}\mathrm{D}_y^{2l-1}f(-1, -1)$$
$$-\theta_k(-x)\theta_l(y)\mathrm{D}_x^{2k-1}\mathrm{D}_y^{2l-1}f(-1, 1) - \theta_k(x)\theta_l(-y)\mathrm{D}_x^{2k-1}\mathrm{D}_y^{2l-1}f(1, -1)$$
$$+ \theta_k(x)\theta_l(y)\mathrm{D}_x^{2k-1}\mathrm{D}_y^{2l-1}f(1, 1)],$$

where $\theta_k(x) = \theta_{k,+}(x)$ (hence $\theta_{k,-}(x) = -\theta_k(-x)$) for $k \in \mathbb{N}$.

This seemingly peculiar function p_r was constructed in a systematic way that will be detailed further on. First, we illustrate the accelerated convergence of modified Fourier series, next we will prove that p_r does indeed satisfy the required properties.
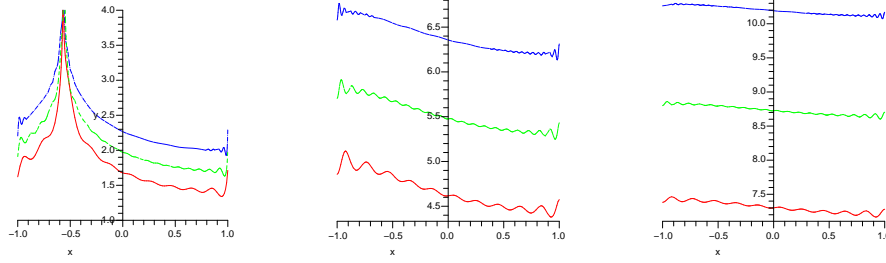
Figure 6: The number of significant decimal digits in approximating $f(x,y) = \sin(1+x-2y)$ along the boundary $y = 1$ by modified Fourier expansions with $r = 0$ (left), $r = 1$ and $r = 2$ (right): solid, dashed and dotted lines represent truncation for $n_1, n_2 \geq N$, where $n = 10, 20, 40$ respectively.

Fig. 5 displays the error committed when approximating $f(x,y) = \sin(1 + x - 2y)$ with 'plain' modified Fourier expansion (top row) and by subtracting p. The improvement in performance is consistent with theory.

In Fig. 6 we present the number of significant digits (that is, $-\log_{10} |\text{error}|$) when the above function is approximated on the boundary $y = 1$. The asymptotic decay appears to be consistent with $\mathcal{O}(n^{-2r-1})$, one less than inside the domain. This can be confirmed using Sheehan Olver's technique of Lerch functions (Olver 2007). Note an important point. While for $r = 0$ (the standard modified Fourier) the difference between convergence inside the domain and on the boundary is $\mathcal{O}(n^{-2})$ vs $\mathcal{O}(n^{-1})$, for $r = 2$ it is $\mathcal{O}(n^{-6})$ compared with $\mathcal{O}(n^{-5})$. The latter is, relatively speaking, of a marginal nature.

Proposition 4 *For every $r \in \mathbb{N}$ and $j_1, j_2 = 0, 1, \ldots, r - 1$ it is true that*

$$\mathrm{D}_x^{2j_1+1}\mathrm{D}_y^{2j_2+1}p_r(x,y) = \mathrm{D}_x^{2j_1+1}\mathrm{D}_y^{2j_2+1}f(x,y), \qquad (x,y) \in \partial[-1,1]^2. \tag{2.12}$$

In addition, for $j = 0, \ldots, r - 1$ and $l = 0, 1, 2, \ldots$ it is true that

$$\mathrm{D}_x^{2j+1}\mathrm{D}_y^{2l}p_r(\pm1,y) = \mathrm{D}_x^{2j+1}\mathrm{D}_y^{2l}f(\pm1,y),$$
$$\mathrm{D}_x^{2l}\mathrm{D}_y^{2j+1}p_r(x,\pm1) = \mathrm{D}_x^{2l}\mathrm{D}_y^{2j+1}f(x,\pm1).$$

Proof By symmetry, it is enough to prove (2.12) only along the face $y = 1$. Since the $\theta_{k,\pm}$s are cardinal polynomials of Birkhoff–Hermite interpolation, it follows that, within the above range of j_1 and j_2,

$$\mathrm{D}_x^{2j_1+1}\mathrm{D}_y^{2j_2+1}p_r(x,1) = \sum_{k=1}^{r}[\theta_{k,-}^{(2j_1+1)}(x)\mathrm{D}_x^{2k-1}\mathrm{D}_y^{2j_2+1}f(-1,1)$$
$$+ \theta_{k,+}^{(2j_1+1)}(x)\mathrm{D}_x^{2k-1}\mathrm{D}_y^{2j_2+1}f(1,1)] + \mathrm{D}_x^{2j_1+1}\mathrm{D}_y^{2j_2+1}f(x,1)$$

$$-\sum_{k=1}^{r}[\theta_{k,-}^{(2j_1+1)}(x)\mathrm{D}_x^{2k-1}\mathrm{D}_y^{2j_2+1}f(-1,1)$$

$$+\theta_{k,+}^{(2j_1+1)}(x)\mathrm{D}_x^{2k-1}\mathrm{D}_y^{2j_2+1}f(1,1)]=\mathrm{D}_x^{2j_1+1}\mathrm{D}_y^{2j_2+1}f(x,1),$$

as required.

The remaining assertion follows by the same reasoning. □

If follows immediately that $\frac{\partial\Delta^j f}{\partial n}=\frac{\partial\Delta^j p_r}{\partial n}$ with $j=0,\ldots,r-1$. Thus, the first r terms in expansion (2.10) vanish for $\tilde{f}=f-p$. We can also establish the full asymptotic expansion of classical Poincaré type for the modified Fourier coefficients.

Theorem 5 *The modified Fourier coefficients of the function $\tilde{f}(x,y)=f(x,y)-p_r(x,y)$, where $r\in\mathbb{Z}_+$ (and $p_0\equiv 0$) obey the asymptotic expansion*

$$\hat{\tilde{f}}_{n_1,0}^{[\alpha_1,0]}\sim(-1)^{n_1+\alpha_1}\sum_{k=r}^{\infty}\frac{(-1)^k}{[\pi\mu_{\alpha_1}(n_1)]^{2k+2}}\int_{-1}^{1}[\mathrm{D}_x^{2k+1}f(1,y)-(-1)^{\alpha_1}\mathrm{D}_x^{2k+1}f(-1,y)]\mathrm{d}y,$$

$$\hat{\tilde{f}}_{0,n_2}^{[0,\alpha_2]}\sim(-1)^{n_2+\alpha_2}\sum_{k=r}^{\infty}\frac{(-1)^k}{[\pi\mu_{\alpha_2}(n_2)]^{2k+2}}\int_{-1}^{1}[\mathrm{D}_y^{2k+1}f(x,1)-(-1)^{\alpha_2}\mathrm{D}_y^{2k+1}f(x,-1)]\mathrm{d}x,$$

$$\hat{\tilde{f}}_{n_1,n_2}^{[\alpha_1,\alpha_2]}\sim(-1)^{n_1+n_2+\alpha_1+\alpha_2}\sum_{k=2r}^{\infty}\frac{(-1)^k}{\pi^{2k+4}}\sum_{\substack{l_1+l_2=k\\l_1,l_2\geq r}}\frac{1}{\mu_{\alpha_1}^{2l_1+2}(n_1)\mu_{\alpha_2}^{2l_2+2}(n_2)}$$

$$\times[\mathrm{D}_x^{2l_1+1}\mathrm{D}_y^{2l_2+1}f(1,1)-(-1)^{\alpha_1}\mathrm{D}_x^{2l_1+1}\mathrm{D}_y^{2l_2+1}f(-1,1)$$

$$-(-1)^{\alpha_2}\mathrm{D}_x^{2l_1+1}\mathrm{D}_y^{2l_2+1}f(1,-1)+(-1)^{\alpha_1+\alpha_2}\mathrm{D}_x^{2l_1+1}\mathrm{D}_y^{2l_2+1}f(-1,-1)]$$

for $n_1,n_2\gg 1$. Here $\alpha_1,\alpha_2\in\{0,1\}$ and

$$\mu_\alpha(n)=\begin{cases}n, & \alpha=0,\\ n-\frac{1}{2}, & \alpha=1.\end{cases}$$

Proof Our point of departure are asymptotic expansions for $r=0$, as proved in (Iserles & Nørsett 2007). By virtue of Proposition 4 we deduce that all the terms for $k\leq r-1$ when $\min\{n_1,n_2\}=0$, $k\leq 2r-1$ otherwise, are nil and all the higher derivatives of $\tilde{f}$ match these of f. □

2.3 d-variate cubes

Bearing in mind how involved are the formulæ for polynomial subtraction in two variables, it might appear hopeless to seek them in higher (and even arbitrary) number of variables. Fortunately, we can present a general recursive formula which produces polynomial-subtraction schemes in $[-1,1]^d$ for any $d\geq 1$. The operative word is "recursive" – writing such schemes explicitly is, obviously, both exceedingly complicated and fairly abstruse.

The main idea is to subtract from the function f boundary effects, one variable at a time. Thus, to obtain the bivariate subtraction (2.11), 'taking away' the first derivative along

$\partial[-1,1]^2$, we let

$$
\begin{aligned}
h^{[0]}(x,y) &= f(x,y), \\
h^{[1]}(x,y) &= h^{[0]}(x,y) - \theta_{1,-}(x)h_x^{[0]}(-1,y) - \theta_{1,+}(x)h_x^{[0]}(1,y), \\
h^{[2]}(x,y) &= h^{[1]}(x,y) - \theta_{1,-}(y)h_y^{[1]}(x,-1) - \theta_{1,+}(y)h_y^{[1]}(x,1).
\end{aligned}
\tag{2.13}
$$

It is trivial to verify directly that the outcome is $h^{[2]} = f - p$, where p has been given by (2.11). However, it is more interesting to prove this going with the grain of the recursive process. Thus, $h_x^{[1]}$, by design, vanishes for $x = \pm 1$ and all $y \in [-1,1]$. Moreover, by the same token, $h_y^{[2]} = 0$ for $y = \pm 1$ – but, by virtue of its construction, $h_x^{[2]} = 0$ for $x = \pm 1$.

This line of reasoning immediately generalises to an arbitrary number of variables. Although general formula can be easily given, it probably suffices to present it for $d = 3$,

$$
\begin{aligned}
h^{[0]}(x,y,z) &= f(x,y,z), \\
h^{[1]}(x,y,z) &= h^{[0]}(x,y,z) - \theta_{1,-}(x)h_x^{[0]}(-1,y,z) - \theta_{1,+}(x)h_x^{[0]}(1,y,z), \\
h^{[2]}(x,y,z) &= h^{[1]}(x,y,z) - \theta_{1,-}(y)h_y^{[1]}(x,-1,z) - \theta_{1,+}(y)h_y^{[1]}(x,1,z), \\
h^{[3]}(x,y,z) &= h^{[2]}(x,y,z) - \theta_{1,-}(z)h_y^{[2]}(x,y,-1) - \theta_{1,+}(z)h_y^{[2]}(x,y,1).
\end{aligned}
$$

Moreover, this framework can be generalised also to higher orders of polynomial subtraction. Thus, in a bivariate setting, (2.13) generalises to

$$
\begin{aligned}
h^{[0]}(x,y) &= f(x,y), \\
h^{[1]}(x,y) &= h^{[0]}(x,y) - \sum_{k=1}^{r}[\theta_{k,-}(x)\mathrm{D}_x^{2k-1}h^{[0]}(-1,y) + \theta_{k,+}(x)\mathrm{D}_x^{2k-1}h^{[0]}(1,y)], \\
h^{[2]}(x,y) &= h^{[1]}(x,y) - \sum_{k=1}^{r}[\theta_{k,-}(y)\mathrm{D}_y^{2k-1}h^{[1]}(x,-1) + \theta_{k,+}(y)\mathrm{D}_y^{2k-1}h^{[1]}(x,1)].
\end{aligned}
$$

Of course, having subtracted the contribution of odd derivative(s) along the boundary of $[-1,1]^d$, we must deal with the $(d-1)$-dimensional functions that we have used there. To obtain true *polynomial* subtraction we must subtract from them, in turn, the contribution of odd derivatives on the boundary – in other words, $(d-2)$-variate functions – and so on. In other words, we need to descend along the faces of $[-1,1]^d$, seen as a complex. All this is straightforward in theory, yet might be fairly messy to implement in practice for large d.

3 The hyperbolic cross

In this section we focus just on cubes $[-1,1]^d$, $d \geq 2$: although our theory is more wide ranging and applies to all multivariate domains where we can use modified Fourier expansion, it is sufficient to understand it in a cube to convey its main flavour.

We examine modified Fourier expansion of the function $f(x) = \mathrm{e}^{\pi(x-y)}$ in the square $[-1,1]^2$. Given $r \in \mathbb{Z}_+$, we restrict our attention to 'pure cosine' terms $c_{m,n}^{[r]} = \hat{f}_{m,n}^{[0,0]}$ with r-fold polynomial subtraction, mentioning in passing that exactly the same arguments are valid
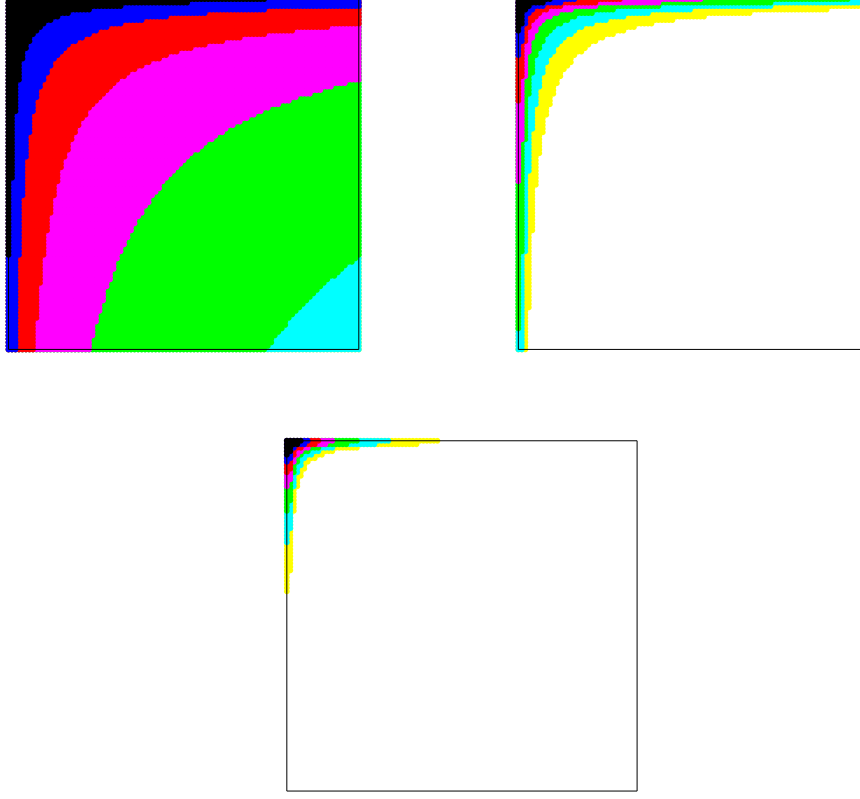
Figure 7: The hyperbolic cross: Contour lines of $|c_{m,n}^{[r]}| > 10^{-k}$ for (in order of decreasing shading) $k = 2, 3, \ldots, 9$ for the function $f(x, y) = e^{\pi(x-y)}$ and $r = 0, 1, 2$. We display $m, n = 0, 1, \ldots, 100$

also in regard to other expansion coefficients. We can calculate explicitly

$$c_{m,n}^{[0]} = \frac{4(-1)^{m+n} \sinh^2 \pi}{\pi^2(1+m^2)(1+n^2)}, \qquad m, n \in \mathbb{Z}_+;$$

$$c_{m,n}^{[1]} = \begin{cases} \dfrac{(\pi^2 - 6)^2 \sinh^2 \pi}{9\pi^2}, & m = n = 0, \\[2mm] \dfrac{2(-1)^m(\pi^2 - 6)\sinh^2 \pi}{3\pi^2 m^2(1+m^2)}, & m \in \mathbb{N}, \; n = 0, \\[2mm] \dfrac{2(-1)^n(\pi^2 - 6)\sinh^2 \pi}{3\pi^2 n^2(1+n^2)}, & m = 0, \; n \in \mathbb{N}, \\[2mm] \dfrac{4(-1)^{m+n} \sinh^2 \pi}{\pi^2 m^2 n^2(1+m^2)(1+n^2)}, & m, n \in \mathbb{N}; \end{cases}$$
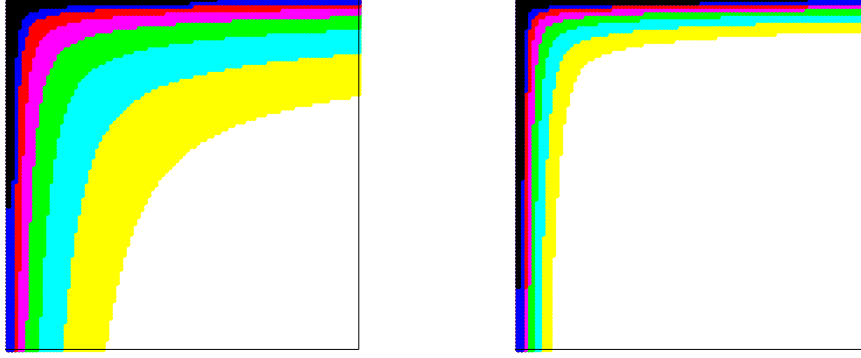
Figure 8: Contour lines of $|d^{[r]}_{m,n}| > 10^{-k}$ for (in order of decreasing shading) $k = 2, 3, \ldots, 9$ for the function $f(x,y) = e^{\pi(x-y)}$ and $r = 1, 2$. We display $m, n = 0, 1, \ldots, 100$

$$c^{[2]}_{m,n} = \begin{cases} -\dfrac{(7\pi^4 - 60\pi^2 + 360)^2 \sinh^2 \pi}{32400\pi^2}, & n = m = 0, \\[3mm] \dfrac{(-1)^m (7\pi^4 - 60\pi^2 + 360) \sinh^2 \pi}{90\pi^2 m^4 (1 + m^2)}, & m \in \mathbb{N},\ n = 0, \\[3mm] \dfrac{(-1)^n (7\pi^4 - 60\pi^2 + 360) \sinh^2 \pi}{90\pi^2 n^4 (1 + n^2)}, & m = 0,\ n \in \mathbb{N}, \\[3mm] \dfrac{4(-1)^{m+n} \sinh^2 \pi}{\pi^2 m^4 n^4 (1 + m^2)(1 + n^2)}, & m, n \in \mathbb{N}. \end{cases}$$

The size of the coefficients is consistent with the theory of Section 2. However, the *pattern* of decay exhibited by the coefficients is also of interest and it is displayed in Fig. 7. We have shaded there the entries corresponding to the coefficients $c^{[r]}_{m,n}$ (with $c^{[r]}_{0,0}$ at the top left corner) according to their magnitude. Thus, $|c^{[r]}_{m,n}| > 10^{-2}$ is painted black and, subsequently, $10^{-k+1} > |c^{[r]}_{m,n}| > 10^{-k}$, for $k = 3, 4, \ldots, 8$, in increasingly lighter shades of grey. Finally, $|c^{[r]}_{m,n}| < 10^{-8}$ is painted white.

Two striking observations are clear from Fig. 7. Although both are fairly obvious by this stage of our exposition and should come as no surprise, they are nonetheless brought home fairly vividly by the 'size plot'. Firstly, the magnitude of the elements decreases very rapidly as r grows. Secondly, elements of similar magnitude are grouped together, with large elements hugging the axes $m = 0$ and $n = 0$ and rapid decay exhibited once one moves diagonally 'into' the lattice (m, n). This is precisely the phenomenon of the *hyperbolic cross,* as identified by Babenko (1960) in the different setting of multivariate Fourier expansions and is precipitated by the fact that $c^{[r]}_{m,n} = \mathcal{O}\big((mn)^{-2r-2}\big)$.

For additional insight, we briefly illustrate the same procedure when subtracting only odd derivatives at the corners. Thus, we choose p to satisfy condition (2.7) and its higher order

generalizations, but nothing more. We have

$$d_{m,n}^{[0]} = c_{m,n}^{[0]}, \qquad m,n \in \mathbb{Z}_+;$$

$$d_{m,n}^{[1]} = \begin{cases} -\dfrac{(\pi^4 - 36)\sinh^2 \pi}{9\pi^2}, & m = n = 0, \\[2ex] \dfrac{2(-1)^m(6 - \pi^2)\sinh^2 \pi}{3\pi^2(1 + m^2)} - \dfrac{2(-1)^m}{3m^2(1 + m^2)}, & m \in \mathbb{N},\ n = 0, \\[2ex] \dfrac{2(-1)^n(6 - \pi^2)\sinh^2 \pi}{3\pi^2(1 + n^2)} - \dfrac{2(-1)^n}{3n^2(1 + n^2)}, & m = 0,\ n \in \mathbb{N}, \\[2ex] -\dfrac{4(-1)^{m+n}(m^2 + n^2 + 1)\sinh^2 \pi}{\pi^2 m^2 n^2(1 + m^2)(1 + n^2)}, & m,n \in \mathbb{N}; \end{cases}$$

$$d_{m,n}^{[2]} = \begin{cases} \dfrac{(7\pi^6 - 30\pi^4 + 1080)\sinh^2 \pi}{270\pi^2}, & n = m = 0, \\[2ex] \dfrac{(-1)^m}{90}\dfrac{\sinh^2 \pi}{1 + m^2}\left[\dfrac{7\pi^4 - 60\pi^2 + 360}{\pi^2} + \dfrac{7\pi^2}{m^2} + \dfrac{60}{m^4}\right], & m \in \mathbb{N},\ n = 0, \\[2ex] \dfrac{(-1)^n}{90}\dfrac{\sinh^2 \pi}{1 + n^2}\left[\dfrac{7\pi^4 - 60\pi^2 + 360}{\pi^2} + \dfrac{7\pi^2}{n^2} + \dfrac{60}{n^4}\right], & m = 0,\ n \in \mathbb{N}, \\[2ex] \dfrac{4(-1)^{m+n}(m^4 + 2m^2 n^2 + n^4)}{\pi^2 m^4 n^4(1 + m^2)(1 + n^2)}\sinh^2 \pi, & m,n \in \mathbb{N}. \end{cases}$$

Fig. 8 clearly demonstrates that interpolation of the normal derivative along the entire boundary is highly advantageous for an accelerated decay of coefficients near the axes, where either m or n is small. Inside the quadrant, for large m and n, acceleration is already quite effective indeed – and this is interpolating just a few partial derivatives at the corners of the domain.

We continue the discussion assuming full subtraction of the normal derivative along the boundary. The obvious manner of truncating the bivariate modified Fourier expansion (2.6) (with or without polynomial subtraction) is by setting to zero all $\widetilde{\widehat{f}}_{m,n}^{[\alpha,\beta]}$ for $m,n \geq N$ for some N, where the size of N scales like an inverse power law of the desired accuracy. In other words, we require an expansion with $(2N+1)^2 = \mathcal{O}(N^2)$ coefficients: such coefficients need be precomputed and, once the expansion has been formed, they all need be taken into account once the approximate value of f is calculated for some (x,y). However, a cursory look at Fig. 7 confirms that *the overwhelming majority of these coefficients are not adding at all to the accuracy of the approximation!* They are as small – if not smaller – as the coefficients we have discarded by choosing our value of N. Thus, if we wish, for example, to discard all $c_{m,n}^{[r]}$s smaller than 10^{-8}, we can disregard all the 'white' terms in Fig. 7 – for $r = 1$ this means that only a small number of terms survives, while for $r = 2$ it is much smaller yet!

Although the hyperbolic cross is not unique to the current setting and is quite well known in the context of Fourier and Chebyshev expansions (Delvos & Schempp 1989, Temlyakov 1993, Wasilkowski & Woźniakowski 1995), it is fairly difficult to implement in a classical Fourier setting. Fourier coefficients are usually derived with (bivariate) FFT, and the latter requires the availability of data in an entire $N \times N$ square lattice and calls for $\mathcal{O}(N^2 \log N)$ operations. Although modern non-standard versions of FFT manage to deal with this situation (Fenn, Kunis & Potts 2006), they are neither easy nor straightforward to implement. However, methods for the evaluation of modified Fourier coefficients (Iserles & Nørsett 2007) (which can be readily extended to cater for polynomial subtraction) obtain them one-by-one. In other

words, provided that we restrict attention to a subset consisting of any M coefficients, the cost of their computation is just $\mathcal{O}(M)$. This provides a means to combine the hyperbolic cross with modified Fourier and polynomial subtraction which is both simple to comprehend and easy to implement.

What applies to $d = 2$ applies with even greater force to larger number of variables. The proportion of points that need be considered in a d-dimensional lattice to attain given accuracy exhibits *very* weak growth as a function of $d \geq 1$.

We formally define the *hyperbolic cross* $K_d(t)$ as the set of all $\boldsymbol{n} \in \mathbb{Z}_+^d$ such that

$$\bar{n}_1 \bar{n}_2 \cdots \bar{n}_d \leq t, \qquad \text{where} \qquad \bar{n}_k = \max\{n_k, 1\}.$$

The number of points $\kappa_d(t)$ in $K_d(t)$ has been calculated in (Dobrovol'skii & Roshchenya 1998).

Proposition 6 *The number of points in the hyperbolic cross $K_d(t)$ is*

$$\kappa_d(t) = \frac{t(\log t)^{d-1}}{(d-1)!} + \textit{lower-order terms.} \tag{3.1}$$

Proof We forego the longer (and more general) proof in (Dobrovol'skii & Roshchenya 1998) for an elementary inductive argument. Clearly, $\kappa_1(t) = t$ and this is consistent with (3.1). For $d \geq 2$ we note the recurrence relation

$$K_d(t) = \sum_{m=1}^{\lfloor t \rfloor} K_{d-1}\left(\frac{t}{m}\right).$$

Therefore, by induction (and retaining only leading terms)

$$
\begin{aligned}
\kappa_d(t) &= \sum_{m=1}^{\lfloor t \rfloor} \kappa_{d-1}\left(\frac{t}{m}\right) \approx \frac{1}{(d-2)!} \sum_{m=1}^{\lfloor t \rfloor} \frac{t}{m}\left(\log \frac{t}{m}\right)^{d-2} \\
&\approx \frac{1}{(d-2)!} \int_1^t \frac{t}{m}\left(\log \frac{t}{m}\right)^{d-2} \mathrm{d}m = \frac{t}{(d-2)!} \int_1^t \frac{(\log x)^{d-2}}{x} \mathrm{d}x \\
&= \frac{t(\log t)^{d-1}}{(d-1)!},
\end{aligned}
$$

as stipulated by (3.1). □

Suppose that, given $f \in \mathrm{C}^\infty([-1,1]^d)$, there exists $\lambda > 0$ such that

$$\|\mathrm{D}_{x_1}^{n_1} \mathrm{D}_{x_2}^{n_2} \cdots \mathrm{D}_{x_d}^{n_d} f\|_{\mathrm{L}_\infty([-1,1]^d)} \leq (\pi\lambda)^{n_1+n_2+\cdots+n_d}, \qquad n_1, \ldots, n_d \in \mathbb{Z}_+. \tag{3.2}$$

(It is possible to embrace more general point of view, similar to that in (Temlyakov 1993), but at present we strive for insight, rather than for maximal generality.)

Similarly to Theorem 5, we have for all $r \geq 0$ and $n_1, n_2 \ldots, n_d \gg 1$

$$c_{n_1,n_2,\ldots,n_d}^{[r]} \sim \frac{(-1)^{|\boldsymbol{n}|+r} \mathcal{S}[f]}{\pi^{2r+2d} n_1^{2r+2} n_2^{2r+2} \cdots n_d^{2r+2}},$$

where $\mathcal{S}[f]$ is a linear combination (with ± 1 coefficients) of $\mathrm{D}_{x_1}^{2r+1}\mathrm{D}_{x_2}^{2r+1}\cdots\mathrm{D}_{x_d}^{2r+1}$ at the 2^d vertices of the d-cube. Disregarding higher-order terms, we thus deduce from (3.2) that

$$|c_{\boldsymbol{n}}^{[r]}| \le \frac{2^d \lambda^{2rd+d}}{(n_1 n_2 \cdots n_d)^{2r+2}}, \qquad n_1, n_2, \ldots, n_d \gg 1. \tag{3.3}$$

Suppose that we agree to discard all coefficients $c_{\boldsymbol{n}}^{[r]}$ less than $\varepsilon > 0$ in magnitude. Then, by (3.3),

$$\boldsymbol{n} \in K_d(t_{d,r,\lambda}), \qquad \text{where} \qquad t_{d,r,\lambda} = \lambda^d \left(\frac{2^d \lambda^{-d}}{\varepsilon} \right)^{\frac{1}{2r+2}}.$$

Now, by (3.1) and assuming for simplicity that $\lambda \ge 1$,

$$\kappa_d(t_{d,r,\lambda}) \approx \frac{2^{\frac{d}{2r+2}} \lambda^{d-\frac{d}{2r+2}}}{(d-1)! \varepsilon^{\frac{1}{2r+2}}} \left[\left(d - \frac{d}{2r+2} \right) \log \lambda + \frac{d}{2r+2} \log 2 + \frac{1}{2r+2} \log \varepsilon^{-1} \right]^{d-1}.$$

Assuming that the ε^{-1} terms dominates, we deduce that we require just about

$$\frac{2^{\frac{d}{2r+2}} \lambda^{d-\frac{d}{2r+2}}}{(d-1)!(2r+2)^{d-1}} \varepsilon^{-\frac{1}{2r+2}} (\log \varepsilon^{-1})^{d-1}$$

terms.

Brief examination of Theorem 5 and the more general (given only for $r = 0$, but all this generalises readily along the lines of that theorem) d-variate expressions in (Iserles & Nørsett 2007), it is easy to verify that, for any $d \ge 2$, we have the following 'layers' of coefficients, corresponding to different faces of the complex $[-1,1]^d$. In each k dimensions, $k = 0, 1, \ldots, d$, we have $2^k \binom{d}{k}$ different types of coefficients $\hat{f}_{\boldsymbol{n}}^{[\boldsymbol{\alpha}]}$, where only k of the n_js are nonzero and the corresponding α_js are either 0 or 1 (if $n_j = 0$ then we also set $\alpha_j = 0$). Since each such coefficient is a k-variate integral, we need to replace d with k in the above estimate of the number of terms. Altogether, we have, after some elementary algebra,

$$1 + \sum_{k=1}^{d} 2^k \binom{d}{k} \frac{2^{\frac{k}{2r+2}} \lambda^{k-\frac{k}{2r+2}}}{(k-1)!(2r+2)^{k-1}} \varepsilon^{-\frac{1}{2r+2}} (\log \varepsilon^{-1})^{k-1}$$

$$= 1 + \varepsilon^{-\frac{1}{2r+2}} 2^{1+\frac{1}{2r+2}} \lambda^{1-\frac{1}{2r+2}} \mathrm{L}_{d-1}^{(1)}(-2^{1+\frac{1}{2r+2}} \lambda^{1-\frac{1}{2r+2}} \log \varepsilon^{-1}), \tag{3.4}$$

where $\mathrm{L}_n^{(\alpha)}$ is the (generalised) Laguerre polynomial (Rainville 1960).

ε	$r = 0$	$r = 1$	$r = 2$	$r = 3$
10^{-06}	116,182	2,623	742	395
10^{-08}	1,530,224	10,897	2,098	921
10^{-10}	18,986,367	42,695	5,596	2,026
10^{-12}	226,705,024	161,062	14,375	4,295

Table 1: The estimate (3.4) of the total number of coefficients $\ge \varepsilon$ for $d = 2$ and $\lambda = 1$.

In Tables 1 and 2 we display the number of terms greater than ϵ in magnitude, as predicted by (3.4), in two and four variables. Of course, (3.4) is a coarse (and grossly pessimistic in

ε	$r = 0$	$r = 1$	$r = 2$	$r = 3$
10^{-06}	37,439,571	622,278	159,081	80,461
10^{-08}	829,237,918	4,310,161	747,652	311,475
10^{-10}	15,532,963,636	25,351,179	2,988,241	1,026,185
10^{-12}	260,907,636,317	134,000,994	10,741,885	3,041,963

Table 2: The estimate (3.4) of the total number of coefficients $\geq \varepsilon$ for $d = 4$ and $\lambda = 1$.

our experience) estimate. Yet, even if, on the face of it, we are dealing here with fairly large numbers, we must also bear in mind the fairly large number of variables! Thus, considering the bottom row in Table 2, even without polynomial subtraction (the first column) this number of terms is equivalent to using a four-dimensional cubical lattice of coefficients (of either conventional or modified Fourier expansion) of side 715, while triple polynomial subtraction, $r = 3$, requires side of length 42. Now, taking 715 coefficients in univariate Fourier expansion of a non-periodic function and incurring an error of $\approx 10^{-12}$ is surely quite remarkable – taking just 42 terms even more so!

It might be too optimistic to hope for the curse of dimensionality to be completely lifted from multivariate approximation. Having said so, the combined magic of modified Fourier expansion, polynomial subtraction and the hyperbolic cross can draw away much of its sting. Clearly, our current understanding of these phenomena and their interaction is far from complete. Yet even partial results and initial numerical experience indicate that they present us with a powerful approach to multivariate approximation.

Acknowledgements

References

Babenko, K. (1960), 'Approximation of periodic functions of many variables by trigonometric polynomials', *Soviet Maths* **1**, 513–516.

Delvos, F.-J. & Schempp, W. (1989), *Boolean Methods in Interpolation and Approximation*, Longman, Harlow.

Dobrovol'skii, N. M. & Roshchenya, A. L. (1998), 'Number of lattice points in the hyperbolic cross', *Mathematical Notes* **63**, 319–324.

Farin, G. (1997), *Curves and Surfaces for CAGD: A Practical Guide*, 4th edn, Academic Press, San Diego.

Fenn, M., Kunis, S. & Potts, D. (2006), 'Fast evaluation of trigonometric polynomials from hyperbolic crosses', *Numerical Algorithms* **41**, 339–352.

Huybrechs, D. & Vandewalle, S. (2006), 'On the evaluation of highly oscillatory integrals by analytic continuation', *SIAM J. Num. Anal.* **44**, 1026–1048.

Iserles, A. & Nørsett, S. P. (2005), 'Efficient quadrature of highly oscillatory integrals using derivatives', *Proc. Royal Soc. A* **461**, 1383–1399.

Iserles, A. & Nørsett, S. P. (2006*a*), From high oscillation to rapid approximation I: Modified Fourier expansions, Technical Report NA2006/05, DAMTP, University of Cambridge.

Iserles, A. & Nørsett, S. P. (2006*b*), From high oscillation to rapid approximation II: Expansions in polyharmonic eigenfunctions, Technical Report NA2006/05, DAMTP, University of Cambridge.

Iserles, A. & Nørsett, S. P. (2007), From high oscillation to rapid approximation III: Multivariate expansions, Technical Report NA2007/01, DAMTP, University of Cambridge.

Kantorovich, L. V. & Krylov, V. I. (1958), *Approximate Methods of Higher Analysis*, 3rd edn, Interscience, New York.

Lorenz, G. G., Jetter, K. & Riemenschneider, S. D. (1983), *Birkhoff Interpolation*, Addison–Wesley, London.

Olver, S. (2006), 'Moment-free numerical integration of highly oscillatory functions', *IMA J. Num. Anal.* **26**, 213–227.

Olver, S. (2007), On the convergence rate of modified Fourier series, Technical report, DAMTP, University of Cambridge, Cambridge. NA2007/02.

Rainville, E. D. (1960), *Special Functions*, Macmillan, New York.

Roache, P. J. (1978), 'A pseudo-spectral FFT technique for non-periodic problems', *J. Comp. Phys.* **27**, 204–220.

Temlyakov, V. (1993), *Approximation of Periodic Functions*, Nova Sci., New York.

Wasilkowski, G. & Woźniakowski, H. (1995), 'Explicit cost bounds of algorithms for multivariate tensor product problems', *J. Complexity* **11**, 1–56.